



University of  
Zurich<sup>UZH</sup>

**USZ** Universitäts  
Spital Zürich

# Adaptive fractionation at the MR-Linac

Master thesis in physics

Yoel Pérez Haas

March 2022

Supervisor: Prof. Dr. Jan Unkelbach  
Tutor: Roman Ludwig

Physik Institut, University of Zurich  
Department of Radiation Oncology, University Hospital Zürich,  
Switzerland

---

## Abstract

Inter-fraction motion of tumors and dose-limiting organs at risk interferes with the treatment quality of radiotherapy. MR guided radiotherapy allows monitoring of inter-fraction motion and direct adjustments of the treatment plan. Whereas standard treatments deliver the same dose in each fraction, adaptive fractionation is an approach to exploit inter-fraction motion by increasing the dose on days when the distance of tumor and organ at risk is large and decreasing the dose on unfavorable days. To evaluate the concept of adaptive fractionation, a dynamic programming algorithm is developed and evaluated on former patients treated at the MR-Linac for abdominal tumors in five fractions to assess a potential tumor dose escalation. Further extensions and adjustments are discussed and evaluated on synthetic patient data to quantify the potential benefit of adaptive fractionation.

On average, adaptive fractionation provided only a small increase in tumor BED. However, individuals with large anatomic variations in between fractions may yield substantial benefits when the optimal patient geometries occur between fraction one and four. Thus, the order of the sparing factors has the biggest impact on the quality of adaptive fractionation. Constraints like minimum and maximum doses per fraction allow to modify the adaptive fractionation plans. The analysis of such constraints only showed a marginal decrease in benefit from adaptive fractionation.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Context and problem</b>	<b>2</b>
2.1	Radiotherapy . . . . .	2
2.2	Fractionation . . . . .	3
2.3	Image guided radiotherapy and adaptive radiotherapy . . . . .	4
2.4	Adaptive fractionation rationale . . . . .	6
<b>3</b>	<b>Methods and materials</b>	<b>8</b>
3.1	Motion quantification . . . . .	8
3.2	Biological effect of adaptive fractionation . . . . .	8
3.3	MDP model of adaptive fractionation . . . . .	9
3.4	Dynamic programming algorithms . . . . .	11
3.4.1	2D algorithms . . . . .	13
3.4.2	3D algorithm . . . . .	15
3.4.3	Code extensions . . . . .	18
3.5	Probability updating . . . . .	19
3.6	Quantification of benefit . . . . .	20
3.7	Patients and treatment plans . . . . .	21
<b>4</b>	<b>Results and evaluations</b>	<b>22</b>
4.1	Evaluated patients . . . . .	22
4.2	Illustration of adaptive fractionation for an example . . . . .	23
4.3	Evaluation of patients . . . . .	28
4.4	Permutation of extreme results . . . . .	32
4.5	Artificial data . . . . .	33
4.6	Tumor dose objective and 3D algorithm . . . . .	34
4.7	Code extensions . . . . .	37
4.7.1	Dose constraint extension . . . . .	37
4.7.2	Risk factor extension . . . . .	38
<b>5</b>	<b>Graphical user interface</b>	<b>40</b>
5.1	Probability distribution menu . . . . .	41
5.2	Extension menu . . . . .	42
5.3	Main input menu . . . . .	43
5.4	Output window . . . . .	44
<b>6</b>	<b>Discussion</b>	<b>45</b>
6.1	Illustration of adaptive fractionation for an example patient . . . . .	45
6.2	Evaluation of patients . . . . .	45
6.3	Permutation of extreme results . . . . .	46
6.4	Artificial data . . . . .	46
6.5	Tumor dose objective and 3D algorithm . . . . .	46
6.6	Algorithm extensions . . . . .	47
<b>7</b>	<b>Conclusion</b>	<b>49</b>
<b>8</b>	<b>Outlook</b>	<b>50</b>
<b>9</b>	<b>Acknowledgements</b>	<b>51</b>

<b>10 Appendix</b>	<b>52</b>
10.1 Code and patient data . . . . .	52
10.2 Tables and Figures . . . . .	52

## 1 Introduction

Radiotherapy treatments aim to kill cancer cells and achieve a tumor control. To do so, ionizing radiation is used to inflict damage at a molecular level and thereby destroy cancer tissue. A downside of radiation treatments is the simultaneous irradiation of healthy tissue close to the tumor. In order to minimize the toxic effects on normal tissue, radiation treatments are usually split up into several sessions, as normal tissue can tolerate higher doses of ionizing radiation if the radiation is split up into fractions [1][2]. Since the tumor and the organs at risk are not stationary, the geometry and distances between them varies from one fraction to another. This interfractional motion worsens the trade-off between tumor coverage and normal tissue sparing. Image guidance technology allows to measure interfractional motion and adapt the treatment plan accordingly.

Adaptive fractionation is a method to exploit inter-fraction motion. The treatment plan is optimized based on the daily patient geometry. Therefore, the dose is increased on favorable treatment days, i.e. when the distance between tumor and dose-limiting organ at risk (OAR) is relatively large and reducing the dose on unfavorable days, i.e. when the tumor and OAR are closer to each other. The adaptation of the dose to the different geometries leads to an improvement of total dose delivered to the OAR versus total dose delivered to the tumor compared to treatment plans, where the same dose is given in each fraction [3][4][5].

Although the idea of adaptive fractionation has been presented previously, clinical translation has been facing substantial hurdles. To implement adaptive fractionation, frequent image updates of the treated region are required to measure and quantify the changes of patient geometry before each irradiation [6]. Cone-beam CTs provided the required information to some degree, but the limited soft tissue contrast limits the potential clinical applications. Magnetic-resonance (MR) scans are well suited as guidance due to their high soft tissue contrast which captures potential tumor motion and extends the range of potential applications to abdominal lesions in proximity to bowel and stomach which exhibit a potential large inter-fraction motion. However, under normal circumstances, acquiring MR-scans for each patient before treatment is not feasible. The MR-Linac enables these daily scans before each irradiation and as a result direct treatment adaptation. In this master thesis an approach for adaptive fractionation at the MR-Linac is presented and evaluated.

To compute an optimal dose for each fraction, different versions of dynamic programming algorithms have been developed. The algorithms propose an optimal dose for each fraction based on the daily patient geometry, dose prescriptions and data from former patients that have been treated at the MR-Linac.

All algorithms have been tested on the data of 16 former MR-Linac patients of the university hospital of Zurich. Additionally, artificially generated patient data with similar geometric variance has been produced to evaluate the adaptive fractionation algorithms on a larger scale. Further, graphical user interfaces have been implemented to facilitate optimal dose calculations even when several parameters can be adjusted.

## 2 Context and problem

Before detailing the methods and implementation of adaptive fractionation, a few key-points of radiotherapy are introduced and the MR-Linac and its operation are described. The resulting problem of adaptive fractionation, that is discussed in this thesis, is then established.

### 2.1 Radiotherapy

In Radiotherapy the aim is to damage and destroy cancer tissue using ionizing radiation. The MR-Linac is equipped with a 6MV linear accelerator which produces X-rays that interact with matter by ionizing molecules within cells. The ejected electrons produced by the interaction are accountable for most of the biological damage as they cause further ionizations in molecules they collide with. This damage is distributed randomly. Whereas direct hits to the DNA are often lethal to the cell, the free radicals produced by the radiation can react with the DNA and therefore also deliver damage indirectly. Since the tumor and the radiation beam can practically never be positioned such that only the tumor is affected by the ionizing radiation, the goal of radiotherapy is to deliver a large dose of ionizing radiation to the tumor while sparing the normal tissue in proximity.[7]

To setup an irradiation plan, in a first step the tumor must be outlined and differentiated from normal tissue. In a standard treatment, computed tomography (CT) scans of the treated volume are taken, which provide anatomical information and density data for radiation dose calculation. If required, additional diagnostic imaging modalities, e.g. magnetic resonance (MR), positron emission tomography (PET) or single photon emission computed tomography (SPECT) scans can be conducted to optimize tumor localization and differentiation from normal tissue with similar electron density to the tumor. The outlined tumor is then expanded to specific target volumes which are embedded into each other. The central volume is the gross tumor volume (GTV) which is the primary tumor mass shown by clinical examination or by imaging. This volume is then extended by the clinical target volume (CTV) which contains additional microscopic diseases that need to be eradicated. A last extension is made with the planning target volume (PTV) which takes setup errors during the future treatment or organ movement and changes in size into account. Since in a standard treatment no additional CT-scans or MR-scans are acquired in between the irradiation session, which can span over several weeks, the PTV margin is essential to cover the uncertainties arising from geometrical variations.[8]

Apart from the target volume, OARs are delineated during treatment planning. The OARs are defined as "those normal tissues which lie adjacent to tumors and may therefore be included within treated volumes, with a risk that the radiation may impair their normal functioning"[8]. The OAR may overlap with the PTV which can significantly constrain the tumor dose prescription as high doses to the tumor would also damage the OAR (see figure 1). Additionally, there can be several organs close to the target volume which lead to multiple dose limitations. In this work only tumors with one dose limiting OAR are considered and analyzed.

With the tumor and the OAR outlined, an optimal dose distribution can be computed. Prescription doses and dose limits give a set of constraints and objectives that should be achieved. An optimization algorithm then computes the optimal achievable dose distribution, given different beam angles and considering the interaction of the ionizing radiation with the tissue.[6][9]

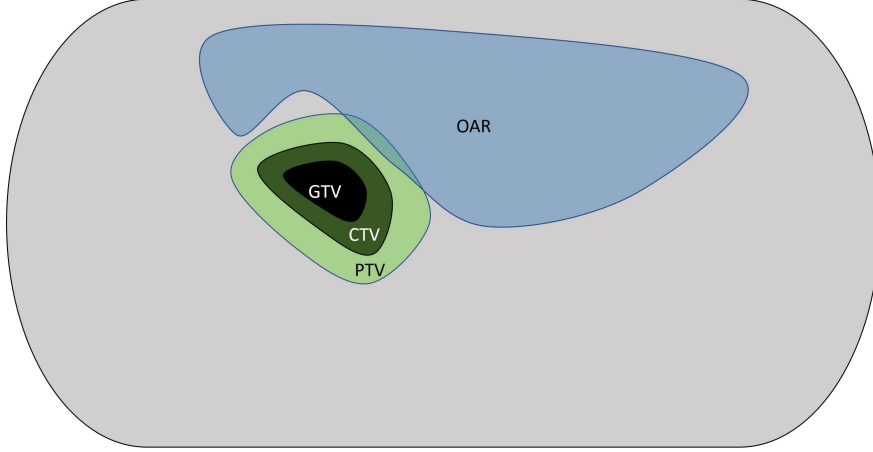


Figure 1: Schematic figure depicting a hypothetical patient geometry. The PTV is in proximity of the OAR and overlaps at some points leading to significant dose constraints.

## 2.2 Fractionation

Fractionation is considered a standard approach to radiotherapy [1][10]. The idea of fractionation is to split the prescribed tumor dose over several fractions and deliver the dose in multiple sessions instead of one single treatment. By fractionating a radiotherapy treatment, a better therapeutic ration, i.e. a better trade-off between tumor control probability and normal tissue complication probability, can be achieved. To cope with the damage induced by ionizing radiation, healthy cells have several repair mechanisms, which are capable of restore the altered DNA. However, cancer cells are often deficient in such repair mechanisms to allow a fast and error prone proliferation. Therefore, the reaction of cancer cells to cell damage is different than in the normal tissue [11]. Splitting the dose to be delivered into several sessions enables normal tissue to repair non-lethal damage in between fractions, while tumor tissue, which is often more sensitive to radiation, does not recover as much from damage as normal tissue.

To model radiosensitivity for fractionated treatments, the linear-quadratic cell survival model is widely used. By assuming that all sublethally damaged cells are completely repaired during the interfractions, the surviving cells are calculated for a n-fractionated treatment. Based on the linear-quadratic cell kill survival model the biological effective dose (BED) can be derived that is used to evaluate the biological effect of fractionation.[12]

The standard BED model is given for a fractionation scheme, where a total physical dose  $D$  is delivered over  $n$  fractions with constant fraction size  $D/n$ :

$$BED = D \left( 1 + \frac{D/n}{\alpha/\beta} \right) \quad (1)$$

The  $\alpha/\beta$  ratio is a measure of the fractionation sensitivity of the cells (tissue specific): cells with a higher  $\alpha/\beta$  ratio are less sensitive to the sparing effect of fractionation [13].

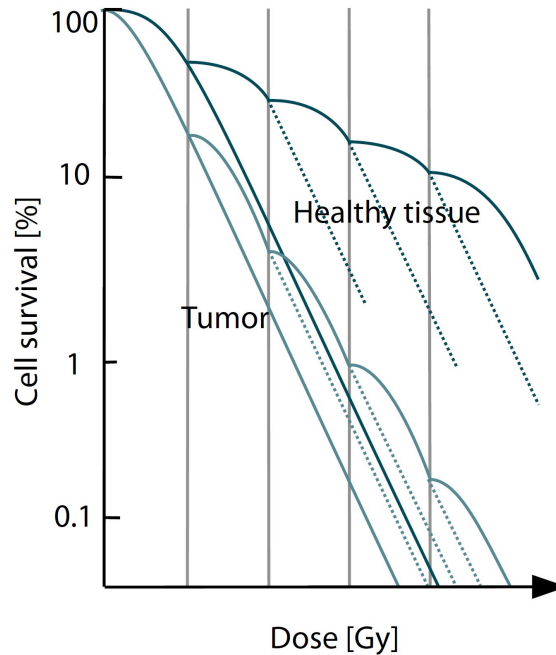


Figure 2: Schematic graph depicting fractionation effects based on the linear quadratic cell survival model. The superior repair capabilities lead to higher cell survival which can be exploited in fractionated treatments. Adapted from *Strahlentherapie und Onkologie* by R. Sauer[14].

A typical fractionated treatment normally consists of 30 to 40 fractions where a constant dose of 1-2Gy are delivered per fraction. A hypofractionated treatment, e.g. stereotactic body radiation therapy (SBRT), can be conducted in five or less fractions and delivers similar amount of doses in total as standard fractionated treatments. Such hypofractionated treatments are superior in specific cases and require high precision [15]. All patients considered in this thesis have undergone SBRT at the MR-Linac and have been treated in only five fractions with high doses around 8Gy to the PTV per fraction.

### 2.3 Image guided radiotherapy and adaptive radiotherapy

To improve the precision of the radiotherapy treatment and minimize the margins around the target volume, due to organ motion as well as setup errors, and to reduce the number of fractions for SBRT, image-guided radiotherapy (IGRT) is applied [16]. Instead of just relying on the information provided by the planning scans, imaging functionalities are integrated into the treatment machines themselves to guide the radiotherapy treatment by providing images before or during each session. IGRT technology can therefore be used to confirm the patient position at the time of each treatment fraction to reduce the safety margins. With these reduced margins, not just hypofractionation is enabled, but also direct adaptation to anatomical or biological changes during the treatment (e.g. due to weight loss, intestinal gas movement or abdominal bloating) are possible with the actual scans of the target area.[17]

IGRT can be conducted with different approaches. First procedures included X-ray images that facilitated and optimized positioning. The technology of image guided therapy has evolved to the application of CT and MR scans as image guidance[18]. MR-guided radiotherapy offers superior soft-tissue contrast for detection of inter-and intrafractional changes [19]. The detected changes can then be directly considered for treatment adaptation. Real-time imaging (cine MRI) further allows the irradiation of moving targets (e.g. lung tumors) with gated dose delivery.



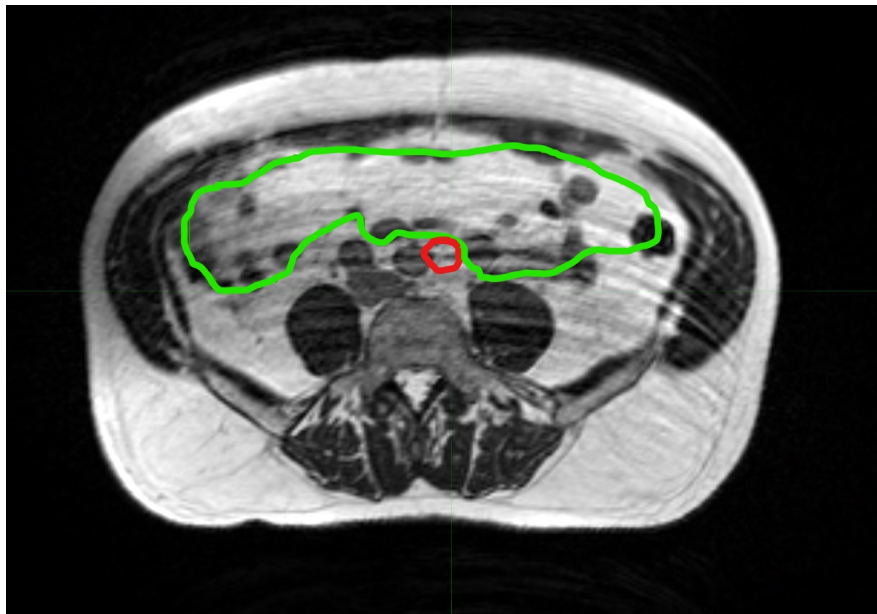
The University Hospital Zürich (USZ) acquired a MR-Linac (MRIdian, Viewray) and has been treating patients with the device since 2019. The system combines a 0.35T MRI with a linear accelerator. As both MRI and linac share the same isocenter, the patient does not need to be moved after imaging.

An integrated treatment planning system allows the creation and adaptation of treatment plans based on MR scans and CT scans. The on-table treatment adaptation registers the daily scans to the primary planning image and initial planning contours. The original treatment plans can then be recalculated on the daily anatomy and corresponding contours.[20]

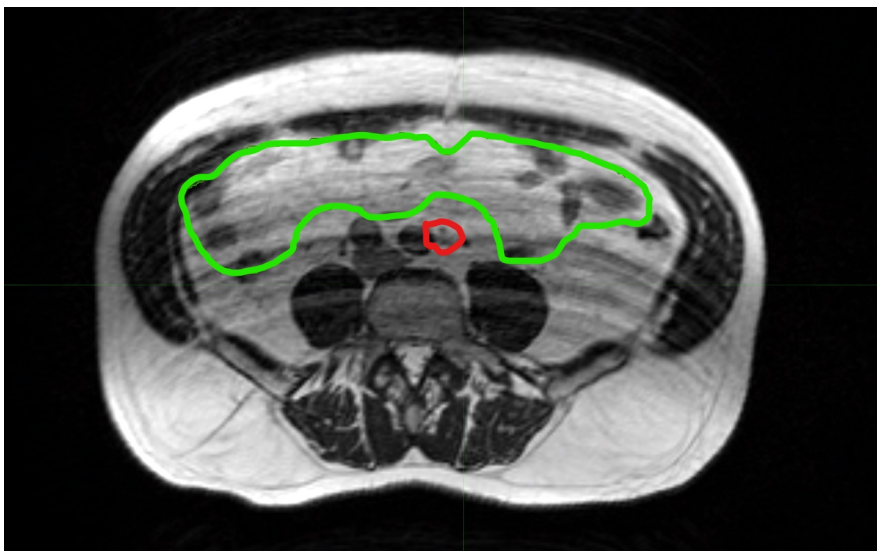
With the precondition set to perform adaptive radiotherapy, the MR-Linac is used to actively adapt treatment plans before the delivery in each fraction. The standard approach at the USZ is to use the updated plans for each fraction, but delivering a uniform dose in each fraction. The fractionation size is therefore not adapted based on the daily anatomy. Thus, the objective of this thesis is to develop and evaluate a procedure to conduct adaptive fractionation at the MR-Linac. A potential application would consist of up- or down-scaling an adapted dose distribution produced by the MR-Linac planning system.

## 2.4 Adaptive fractionation rationale

As the patient geometry is not static over the course of a radiotherapy treatment, the adaptation of the treatment plan can lead to significant benefits in OAR sparing and tumor dose escalation. Especially, patients treated for abdominal tumors can stongly benefit from IGRT due to the potential high interfractional motion [21]. Figure 3 illustrates interfractional motion captured at the MR-linac. The distance between the tumor and the OAR varies significantly.



(a)



(b)

Figure 3: Example of interfractional motion. These scans were acquired on two different days during a five fraction treatment. The slices show an obvious difference in distance between the abdominal tumor (outlined in red) and the OAR which is the bowel (outlined in green)

The dose distribution has been adapted accordingly for both fractions at the time of treatment while the fraction size was chosen to be the same in both days. Taking into account the relatively large distance between tumor and OAR on the second day, one could consider delivering a lower

dose on day one and a larger dose on day two, thereby taking advantage of the better OAR sparing. If the prescribed tumor dose is constrained by the maximum dose to be delivered to the OAR, the exploitation of more favourable days, where the distance between tumor and OAR is large, can lead to tumor dose escalation while staying below the OAR maximum dose constraint. If most of the treatment days consist of good patient geometries, an adaptation of fraction size may further reduce the dose accumulated in the OAR.

A posteriori, when scans for all fractions are present, computing optimal doses for each fraction to maximize tumor dose can be done by solving an optimization problem. In reality, we do not have exact information about future anatomy geometries. Thus, the decision on what dose to deliver in each fraction to maximize tumor dose is non trivial.

With the starting point, where adaptive radiotherapy plans are available from the MR-linac, this work was intended to contribute in the following points:

1. To compute the optimal dose to deliver in each fraction, knowing today's geometry and the dose delivered in previous fractions, but not knowing the patient geometry in future fractions, different dynamic programming algorithms have been developed - extending the work of Ramakrishnan et al.[22].
2. The algorithms were tested on patients previously treated at the MR-Linac with 5-fraction SBRT for abdominal lesions near bowel, stomach, or duodenum. Thus, the potential benefit of adaptive fractionation was estimated for real patient data, extending previous work that only conceptually introduced adaptive fractionation based on synthetic data.
3. To make adaptive fractionation and the developed codes more accessible, a graphical user interface has been implemented and provided on GitHub.

### 3 Methods and materials

In the following subsections, the general approach to adaptive fractionation is discussed. Firstly, the information about organ motion and the biological effect of ionizing radiation are modeled to be integrated into an optimization problem. Secondly, the mathematical framework for the optimization algorithm is set up, followed by three different algorithms given in section 3.4, which were developed to compute the optimal doses. The algorithms are all based on a dynamic programming method but aim at different objectives.

#### 3.1 Motion quantification

MRI scans and dose distributions of the treatment area provide a large amount of information. To compress the information to be better suited for a reinforcement learning algorithm, the interfractional motion is quantified with sparing factors  $\delta$ :

$$\delta_t = \frac{d_t^N}{d_t} \quad (2)$$

where  $d_t^N$  denotes the dose received by the dose-limiting OAR in fraction  $t$  and  $d_t$  the dose received by the tumor. The dose to the OAR  $d_t^N$  is defined by the dose exceeded in 1cc of the OAR ( $D_{1cc}$ ) while the tumor dose  $d_t$  is defined by the dose exceeded in 95% of the PTV volume ( $D_{95\%}$ ). Thus, each fraction is described by a sparing factor where low sparing factors indicate more favourable days with large distances between tumor and OAR. We further assume that inter-fraction motion is random and is described by a Gaussian distribution over  $\delta$  with a patient specific mean  $\mu$  and standard deviation  $\sigma$ .

$$\delta_t \sim \mathcal{N}(\mu, \sigma^2) \quad (3)$$

Similarly, a sparing factor can be set up by taking the ratio of the dose exceeded in 1cc of the OAR and the dose exceeded in 95% of the GTV volume rather than the PTV. To compare and assess the results of adaptive fractionation, the analysis of the extracted patients is conducted on both PTV and GTV based sparing factors.

#### 3.2 Biological effect of adaptive fractionation

The fractionation effect is modeled with the BED model, which is commonly used for isoeffective dose fractionation calculations [23]. It is assumed, that the classic BED model can be extended to varying doses per fraction such that the cumulative BED at the end of treatment is given by the sum of the BED values delivered in individual fractions. Thus, the cumulative BED delivered to the tumor is given by:

$$B_T = \sum_{\tau=1}^t \left( d_\tau + \frac{d_\tau^2}{(\alpha/\beta)_T} \right) \quad (4)$$

where  $d_\tau$  denotes the dose delivered to the tumor in fraction  $\tau$  and  $t$  denotes the number of the last delivered fraction. Consequently, the cumulative BED received by the OAR is:

$$B_N = \sum_{\tau=1}^t \left( \delta_\tau d_\tau + \frac{\delta_\tau^2 d_\tau^2}{(\alpha/\beta)_N} \right) \quad (5)$$

In order to scale the tumor dose to the corresponding dose delivered to the OAR, the sparing factor  $\delta_\tau$  in fraction  $\tau$  is included in the equation. In this thesis, the ratios for the OARs and the tumors are the same for all patients and set to  $(\frac{\alpha}{\beta})_N = 3$  and  $(\frac{\alpha}{\beta})_T = 10$ .

### 3.3 MDP model of adaptive fractionation

The goal of adaptive fractionation is to optimally decide on the doses  $d_t$  that are to be delivered to the tumor in each fraction. The choice of an optimal dose depends on the objective that is to be accomplished. The clinical interest, to deliver a large amount of dose to the tumor, while sparing the organs at risk, leads to three different objectives that were used to set up Markov decision processes (MDP) models which are applied to describe the problem and find a potential solution:

1. To treat a tumor where the desired prescription dose cannot be reached as the OAR is too close to the tumor, the goal was set to maximize the cumulative BED delivered to the tumor subject to the constraint on the cumulative OAR BED.
2. In a case where tumor and OAR are farther apart, the prescribed tumor dose can be obtained without risking the overdosage of the OAR. Therefore, the goal is to minimize the cumulative BED delivered to the OAR subject to delivering the prescribed dose to the tumor.
3. Deciding on which algorithm to use is once again a nontrivial problem at the beginning of a treatment. Hence, an objective has been set, where the goal is to reach the prescribed tumor dose subject to the constraint on the cumulative OAR BED. If the prescribed dose can be reached, the objective is to minimize OAR BED. If the prescribed dose can not be reached, the tumor dose should be maximized.

By acquiring a MR-scan in each fraction and adjusting the dose distribution to the new anatomy, the sparing factor  $\delta_t$  for each fraction  $t$  can be determined. In addition, we know the accumulated BED that has been delivered to the tumor and the OAR in previous fractions. The decision, what dose to deliver in the current fraction, must be taken based on the given information. The difficulty arises from not knowing, whether the remaining future fractions will have favorable or unfavorable patient geometries. The sparing factors, i.e. the interfractional motion, are random variables where the exact future value is unknown.

MDPs are mathematical frameworks to model decision making in discrete, stochastic, sequential environments [24]. A MDP relies on the notion of states, describing the current situation of the agent (the decision-maker), actions or decisions which affect the dynamics of the process and rewards that are observed for each transition between states [25]. Further, the transitions between the states are not fully deterministic, instead they are influenced by probabilistic components. Adaptive fractionation is well suited for such an approach, as each fraction can be described by a state, where a decision must be made.

The adaptive fractionation problem can be formulated as an MDP, illustrated in figure 4, and then be solved by a reinforcement learning approach. Here we first describe the MDP model for a known probability distribution  $P(\delta_t)$  of the sparing factors and afterwards we discuss how to estimate and update the probability distribution. In this application, the notions for an algorithm that maximizes the tumor dose while delivering the maximum OAR dose  $B_{max}^N$ , the first objective introduced above, are given by:

**State:** The state of a patient's treatment is described in each fraction by three values, the fraction number  $t$  and a tuple  $s = (\delta, B^N)$  that specifies today's sparing factor  $\delta$  and the cumulative BED of the OAR that has been delivered so far in previous fractions. Thus, the state of a treatment in fraction  $t$  for a patient with previous sparing factors  $\{\delta_\tau\}_{\tau=1}^t$  treated with doses  $\{d_\tau\}_{\tau=1}^{t-1}$  is

$$s_t = \left( \delta_t, \sum_{\tau=1}^{t-1} \left( \delta_\tau d_\tau + \frac{\delta_\tau^2 d_\tau^2}{(\alpha/\beta)_N} \right) \right) \quad (6)$$

**Action and policy:** The actions correspond to the physical doses  $d_t$  that are delivered to the tumor in a fraction. Thus, a policy specifies for each fraction  $t$  and possible state of the treatment the dose that should be delivered in this state. The doses to be delivered were discretized in 0.1Gy steps since not all possible continuous actions can be considered.

**State transition:** If in fraction  $t$ , the treatment is in state  $s_t = (\delta_t, B)$  and a dose  $d_t$  is delivered to the tumor, the state transitions to

$$s_{t+1} = \left( \delta_{t+1}, B + \delta_t d_t + \frac{\delta_t^2 d_t^2}{(\alpha/\beta)_N} \right) \quad (7)$$

in fraction  $t + 1$ . The BED-component of the future state is calculated by adding the OAR BED delivered in fraction  $t$  to the previously delivered BED  $B$ , which is assumed deterministic (i.e. we do not consider uncertainty in the dose delivery). The sparing factor in fraction  $t + 1$  is random, making the state transition probabilistic. The probability distribution for the state transition is simply given by the probability distribution over the sparing factors,  $P(\delta)$ .

**Reward:** In each fraction  $t$ , the immediate reward  $r_t$  is given by the numeric value of the BED delivered to the tumor in that fraction:

$$r_t = d_t + \frac{d_t^2}{(\alpha/\beta)_T} \quad (8)$$

To account for the cumulative BED constraint in the dose-limiting OAR, the BED must be below  $B_N^{max}$ . To enforce this, a terminal reward of  $-\infty$  is assigned to all terminal states in which the cumulative OAR BED delivered after the last fraction exceeds the constraint value.

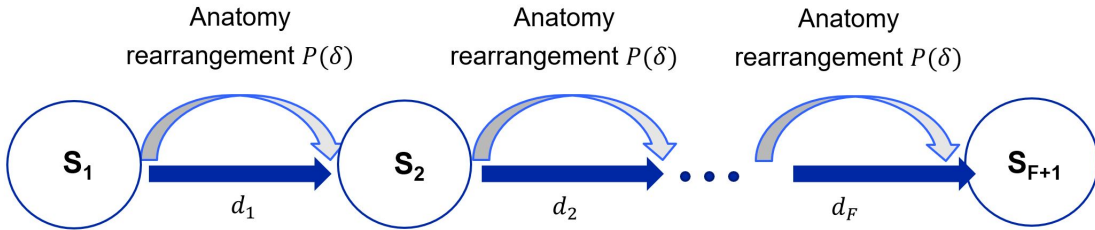


Figure 4: Illustration of the applied MDP. Each state  $s_t$  can be seen as one fraction of a F-fraction treatment that is defined by the accumulated BED and the observed sparing factor. The anatomy rearrangement is a random process while the applied dose  $d$  is the action that will lead to a state transition and reward.

In figure 4 the illustration depicts the state transitions between each fraction. Starting from the first fraction, which is described by a state  $s_1(\delta_1, B)$  where there is no accumulated OAR BED ( $B = 0$ ) and a sparing factor  $\delta_1$  is observed, a state transition is executed by delivering a dose  $d_1$  (the action). The subsequent state  $s_2(\delta_2, B)$  has a probabilistic component since the sparing factor  $\delta_2$  can not be predicted in the first state and is assumed to be independent from the action  $d_1$ . In contrast, the accumulated OAR BED is computed by applying equation (5). Furthermore, the reward is also dependent on the action that was realized in the first fraction. Continuing from state  $s_2$ , a state transition is realized again by the delivery of a dose  $d_2$ . The resulting state  $s_3$  which describes the starting point of the third fraction can be fully characterized by its sparing factor  $\delta_3$  and the accumulated OAR BED which is only dependent on the state  $s_2$  and the action  $d_2$ . As a result, each fraction can be used as a starting point to compute an optimal dose if the current sparing factor  $\delta$  and the accumulated OAR BED  $B$  are known without further knowledge how this state has been achieved.

The MDPs for objective two and three are set up similarly. All major adaptations to the MDP will be described in the following section.

### 3.4 Dynamic programming algorithms

To solve the MDP, a dynamic programming (DP) algorithm was developed following a similar approach as Ramakrishnan et al.[22]. The DP computes the optimal policies when a perfect model of the environment is given [26]. Within the scope of this thesis, three different algorithms have been produced, where each algorithm handles one of the three objectives introduced in section 3.3. In a first step, the general concept of DP aimed to solve adaptive fractionation is introduced based on the first objective. Then the three different algorithms are described, which are grouped in two 2 dimensional and one 3 dimensional approach.

DP algorithms rely on value functions, which describe the desirability of a state  $s$ . In our application, the value function  $v_t(\delta, B)$  is dependent on the sparing factor  $\delta$  and the accumulated BED  $B$  and contains the information whether an action should be taken to reach that state. The rewards in this application are chosen such that the value provides information about the expected cumulative BED that will be delivered to the tumor. For example, the value function for a specific sparing factor  $\delta_1$  in the first fraction gives the expected value of the total BED that can be delivered to the tumor starting from that state.

In general, the Bellman equation relates the value function in fraction  $t$  to the optimal policy  $\pi$  and the value function in the subsequent fraction. For this application it is given as follows:

$$v_t(\delta, B) = \max_d \left[ d + \frac{d^2}{(\alpha/\beta)_T} + \sum_{\delta'} P(\delta') v_{t+1}(\delta', B + \delta d + \frac{\delta^2 d^2}{(\alpha/\beta)_N}) \right] \quad (9)$$

and

$$\pi_t(\delta, B) = \operatorname{argmax}_d \left[ d + \frac{d^2}{(\alpha/\beta)_T} + \sum_{\delta'} P(\delta') v_{t+1}(\delta', B + \delta d + \frac{\delta^2 d^2}{(\alpha/\beta)_N}) \right] \quad (10)$$

To compute the value for any given state  $s_t(\delta, B)$ , we must not just consider the reward, but also the desirability of the reached state when applying a specific dose  $d$ . In equation (9) it is shown that  $v_t$  can only be calculated if the exact value of  $v_{t+1}$  is already known, which is the value of the next fraction. Similarly, the optimal policy  $\pi_t$  depends on the numeric value of the value function in the subsequent state  $v_{t+1}$ . As a consequence, the value functions and the optimal policies must be calculated iteratively in backward recursion starting from the last fraction. The hypothetical value function  $v_{F+1}$  corresponds to the terminal reward at the end of the treatment after all  $F$  fractions are delivered and is initialized to:

$$v_{F+1}(\delta, B) = \begin{cases} -\infty & (B > B_{max}^N) \\ 0 & (B \leq B_{max}^N) \end{cases}$$

Given that the state transition is not deterministic in the dimension of the sparing factor, the probability distribution of future sparing factors  $P(\delta)$  is included in both Bellman equations (9) and (10) as described in equation (6).

The optimal policy is found by discretizing both actions and states. A discretization is needed, since policies and values can only be calculated for explicit states, i.e. for a specific accumulated BED  $B$  and a specific sparing factor  $\delta$ . As the discretization leads to artifacts in the optimal policy (see figure 5), two adjustments are included in the algorithm.

1. The value function is interpolated in the BED-component of the state. Since equation (9) is monotonically decreasing in the BED-component, the interpolation should not distort the results. The discrete BED states before interpolation had a spacing of 1Gy BED. The spacing was chosen, such that the resulting values and policies were equal to a discretization with a spacing of 0.1Gy BED, i.e. the results were equivalent to a more precise discretization that was then interpolated.
2. One can exploit that in the optimal policy, the last fraction will simply deliver the maximum residual BED to the OAR to maximize the dose to be delivered to the tumor and end up at  $B_{max}^N$ . This second adjustment can be used to directly initialize the value function and optimal policy in the last fraction  $F$  using continuous values of  $d_F$  instead of the discretized actions. Thereby artifacts arising from the discretized actions can be avoided which lead to final states, where the maximum OAR BED  $B_{max}^N$  are not exactly reached.

By interpolating the BED component, not only do the artifacts disappear, but also the computation time is lowered drastically. DP algorithms must check the value and policy of every state. If we consider an OAR limit  $B_{max}^N = 90\text{Gy BED}_3$ , then any dosage between 0 and 90 has to be checked. Therefore, increasing the discretization spacing from 0.1 to 1, reduces the number of calculations by 100 without a loss of quality in the results. An interpolation of the sparing factors was not conducted at this point, but can be considered to reduce computation time.

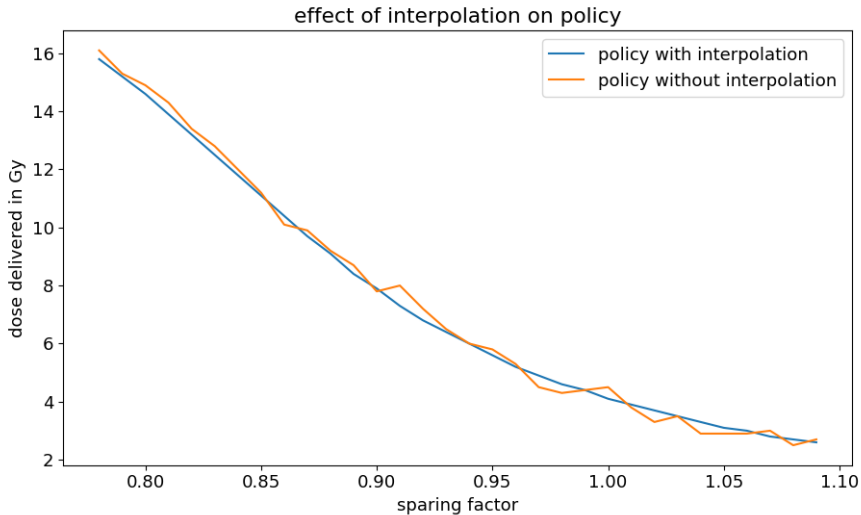


Figure 5: Example of the impact of BED interpolation on the policy of the first fraction for an arbitrary probability distribution. The curve without interpolation proposes higher doses for higher sparing factors which should not happen for a monotonically decreasing value function. The origin of this artifact is the rounding of the future BED for a given dose  $d$ . Which can lead to higher or lower future values that do not correspond to the true future value.

An illustration of the value function and the optimal policy for an arbitrarily chosen probability distribution with  $P(\delta) \sim \mathcal{N}(0.8, 0.1)$  is given in figure 6. The dependencies on the sparing factors and the accumulated OAR BED in equations (9) and (10) can be clearly observed since lower accumulated BEDs are associated with better states and result in high state values as depicted in 6a, while high accumulated doses lead to a lower state value as the limit is nearly reached. A clear impact of the observed sparing factors can also be seen, since the values, the expected tumor BED to be delivered starting from the actual state, reaches the highest values with the



lowest sparing factors. Similarly, the policy in 6b gives an impression of the decision making for any given sparing factor and accumulated BED in the OAR.

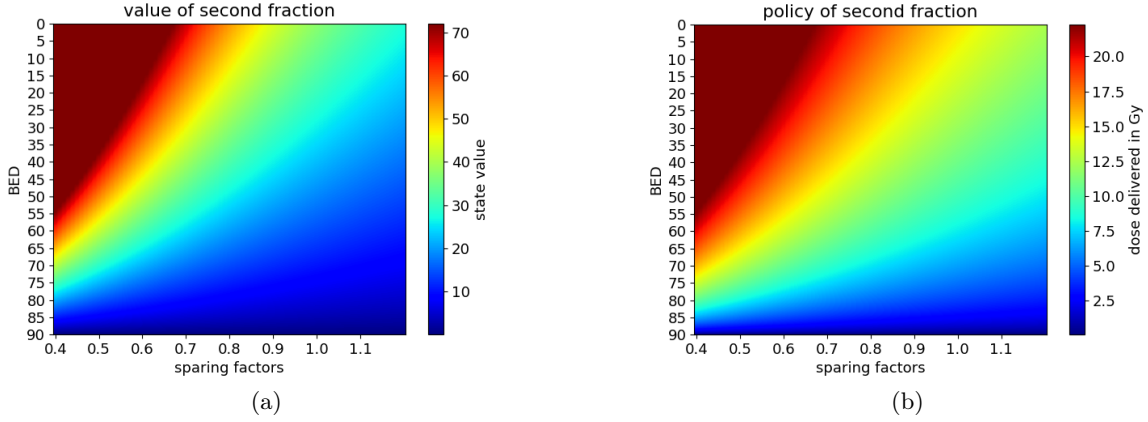


Figure 6: a) Illustration of a value function given a normal distribution. b) Illustration of a policy function based on a normal distribution

### 3.4.1 2D algorithms

The 2D algorithms are defined by their state space which is two dimensional, i.e. depending on two variables, the sparing factor  $\delta$  and the accumulated BED  $B$ . Based on a two dimensional state space, two of the specified objectives can be pursued.

The first objective is to maximize tumor dose subject to the  $B_{max}^N$ . This means the MDP model is defined as the first introduced objective in section 3.3. Thus, the reward is given by the dose delivered to the tumor and the value function represents the expected cumulative BED that can be delivered to the tumor in the remaining fractions, starting from that state and acting according to the optimal policy.

---

#### Algorithm 1 Tumor BED maximization

---

```

for t = F to 1 do                                ▷ Loop through all fractions starting from the last
  if t = F then                                    ▷ Initialization
     $B_{res} \leftarrow B_{max}^N - B^N$                 ▷ residual BED given by difference to the maximum OAR BED
     $d_{res} \leftarrow d | \text{BED}_3(d, \delta) = B_{res}$     ▷ respective physical dose to deliver  $B_{res}$  to OAR
     $V(t, \delta, B^N) \leftarrow \text{BED}_{10}(d_{res})$     ▷ The value is only given by the reward
     $\pi(t, \delta, B^N) \leftarrow d_{res}$ 
  else
     $V(t, \delta, B^N) \leftarrow \max_d \left[ \sum_{\delta} P(\delta) \cdot V(t+1, \delta, B^N + \text{BED}_3(d, \delta)) + \text{BED}_{10}(d) \right]$ 
     $\pi(t, \delta, B^N) \leftarrow \arg\max_d \left[ \sum_{\delta} P(\delta) \cdot V(t+1, \delta, B^N + \text{BED}_3(d, \delta)) + \text{BED}_{10}(d) \right]$ 
  return  $V, \pi$                                     ▷ The policies and the Values for all states are returned
    
```

---

A schematic overview of the 2D algorithm for tumor dose maximization is illustrated in the code block 1. To compute the optimal values for each fraction, the algorithm will loop through all relevant sparing factors  $\delta$  and all possible OAR BEDs  $B^N$ . The computation starts at the last fraction  $F$ , where the optimal dose to be delivered is strictly given by the residual dose that can be delivered to the OAR  $B_{res}$  to precisely reach the OAR limit  $B_{max}^N$ . Based on the residual dose, the respective physical dose to be delivered to the tumor  $d_{res}$  is computed based on equation (5) depicted by  $\text{BED}_{10}(d, \delta)$ . Once the residual physical dose is known, the value for all states

in fraction  $F$  can be computed by assigning the reward (equation 8) since there are no more future states. Similarly, the optimal policy is given by  $d_{res}$ . For all other fractions  $t$ , the value function is not just given by the reward  $BED_{10}(d)$ , but we also have to consider the value of future states that are reached by an action  $d$ . Since the delivery of the dose  $d$  can result in any future BED value, the value function is interpolated (not described in the code block). The matrix  $V(t, \delta, B^N)$  only stores values for discrete  $B^N$  steps which would lead to artifacts in the policy as described in figure 5, if only the discrete BED values would be considered by rounding the future BED values. In the same manner the policy for all states of fraction  $t$  are computed. Vector operations allow to bypass the additional loop through all sparing factors which reduces the computation time, which is not illustrated inside the code block for simplicity. The implemented code also does not loop through all possible OAR BEDs  $B^N$  in fraction  $t = 1$  since at the very beginning, the OAR BED is set to zero.

A drawback from such an algorithm is a potentially overdosed tumor, as the algorithm does not stop at a prescribed dose, since the accumulated tumor dose is not tracked and considered for a decision. Therefore, another approach is to aim at the prescribed tumor dose  $B_{pre}^T$  while minimizing the dose delivered to the OAR. The algorithm then always satisfies the prescribed tumor dose, but may overdose the OAR, pursuing the second objective. Such an algorithm would be better suited when the distance between OAR and tumor is rather large in most fractions. The fundamental change of this second 2D algorithm is that state is now dependent on the sparing factor and the accumulated tumor BED, i.e.  $s(\delta_t, B^T)$  and a new reward definition. The whole set of MDP functions are given by:

$$s_t = \left( \delta_t, \sum_{\tau=1}^{t-1} \left( d_\tau + \frac{d_\tau^2}{(\alpha/\beta)_T} \right) \right) \quad (11)$$

$$s_{t+1} = \left( \delta_{t+1}, B + d_t + \frac{d_t^2}{(\alpha/\beta)_T} \right) \quad (12)$$

$$r_t = - \left( \delta_t d_t + \frac{\delta_t^2 d_t^2}{(\alpha/\beta)_N} \right) \quad (13)$$

The respective rewards for each action are negative, thus it is considered a penalty to deliver dose to the OAR and the value function now represents the expected cumulative BED that will be delivered to the OAR in the remaining fractions, starting from that state and acting according to the optimal policy. Therefore, the algorithm now minimizes the dose to be delivered to the OAR while delivering precisely the prescribed tumor dose. Another adaptation must be made at the initialization of the value function to deliver the remaining dose to reach the prescribed tumor dose:

$$v_{F+1}(\delta, B) = \begin{cases} -\infty & (B \neq B_{pre}^T) \\ 0 & (B = B_{pre}^T) \end{cases}$$

The adaptation of the state dependency, the reward and the initialization are the only modifications as shown in code block 2, where the computation of the first fraction was limited to the case of  $B^N = 0$  similar to the previous code.

---

**Algorithm 2** OAR BED minimization

---

```

for t = F to 1 do                                ▷ Loop through all fractions starting from the last
  if t = F then                                     ▷ Initialization
     $B_{res} \leftarrow B_{pre}^T - B^T$                     ▷ residual BED given by difference to reach  $B_{pre}^T$ 
     $d_{res} \leftarrow d | \text{BED}_{10}(d) = B_{res}$           ▷ respective physical dose to deliver  $B_{res}$  to tumor
     $V(t, \delta, B^N) \leftarrow -\text{BED}_3(d_{res}, \delta)$       ▷ penalty from dose delivered to OAR
     $\pi(t, \delta, B^T) \leftarrow d_{res}$                 ▷ Optimal policy is given by the residual dose
  else
     $V(t, \delta, B^T) \leftarrow \max_d [\sum_{\delta} P(\delta) \cdot V(t+1, \delta, B^T + \text{BED}_{10}(d)) - \text{BED}_3(d, \delta)]$ 
     $\pi(t, \delta, B^T) \leftarrow \text{argmax}_d [\sum_{\delta} P(\delta) \cdot V(t+1, \delta, B^T + \text{BED}_{10}(d)) - \text{BED}_3(d, \delta)]$ 
  return  $V, \pi$                                      ▷ The policies and the Values for all states are returned

```

---

Both of the 2D algorithms allow a fast calculation of five fraction treatment, given the probability distribution. A whole treatment plan can be optimized in a few seconds thus making the algorithms well suited for an online treatment adaptation.

### 3.4.2 3D algorithm

The application of a 2D algorithm requires the determination whether a patient will have several treatment days where the OAR and the tumor are in close proximity and therefore the prescribed dose may not be applied or whether the patient will surely reach the prescribed dose and hence a OAR dose minimization is of interest. These two objectives can be combined to an algorithm that seeks to reach the prescribed tumor  $B_{pre}^T$  dose subject to the maximum OAR dose  $B_{max}^N$  while minimizing the OAR BED if the prescribed dose can be reached as given by the third objective.

Combining the two objectives leads to a MDP structure that is now defined by a state that is dependent on three variables: The sparing factor  $\delta$ , the OAR BED  $B^N$  and the tumor BED  $B^T$  hence its name 3D algorithm. The resulting notions are given below.

$$s_t = \left( \delta_t, \sum_{\tau=1}^{t-1} (\delta_{\tau} d_{\tau} + \frac{\delta_{\tau}^2 d_{\tau}^2}{(\alpha/\beta)_N}), \sum_{\tau=1}^{t-1} (d_{\tau} + \frac{d_{\tau}^2}{(\alpha/\beta)_T}) \right) \quad (14)$$

$$s_{t+1} = \left( \delta_{t+1}, B^N + \delta_t d_t + \frac{\delta_t^2 d_t^2}{(\alpha/\beta)_N}, B^T + d_t + \frac{d_t^2}{(\alpha/\beta)_T} \right) \quad (15)$$

The reward would still be shaped as a penalty for each fraction. But since the OAR BED is be tracked, the penalty does not have to be computed for each fraction but can be integrated into the computation of the final state value. Furthermore, the initialisation of the algorithm is chosen such that the prescribed tumor dose nor the maximum OAR BED are surpassed. Thus, the penalty applied after the last fraction is given by:

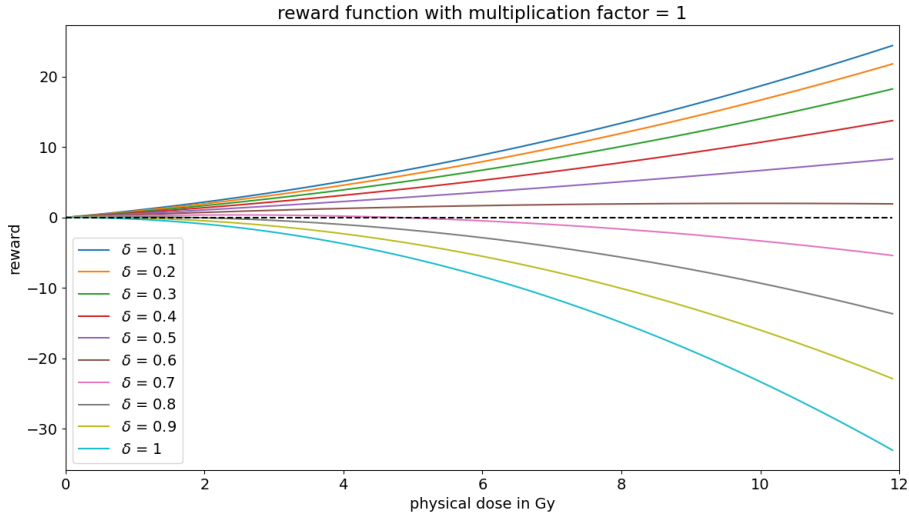
$$v_{F+1}(\delta, B^T, B^N) = (B^T - B_{pre}^T) \cdot f - B^N \quad (16)$$

The final reward/penalty in equation (16) penalizes whenever the tumor is underdosed i.e. when the difference between the prescribed tumor dose and the accumulated tumor BED after the treatment is not zero. The difference is then multiplied by a factor  $f$  to assure that the penalty of underdosing the tumor is weighted more than sparing the OAR. Also, the total accumulated BED in the OAR is subtracted to penalize any dose delivered to the OAR. Altering the multiplication factor sets a limit at which sparing factor OAR sparing becomes more important than delivering dose to the tumor. This becomes clear when the difference between the OAR penalty and the tumor underdosing penalty is considered for a single fraction:

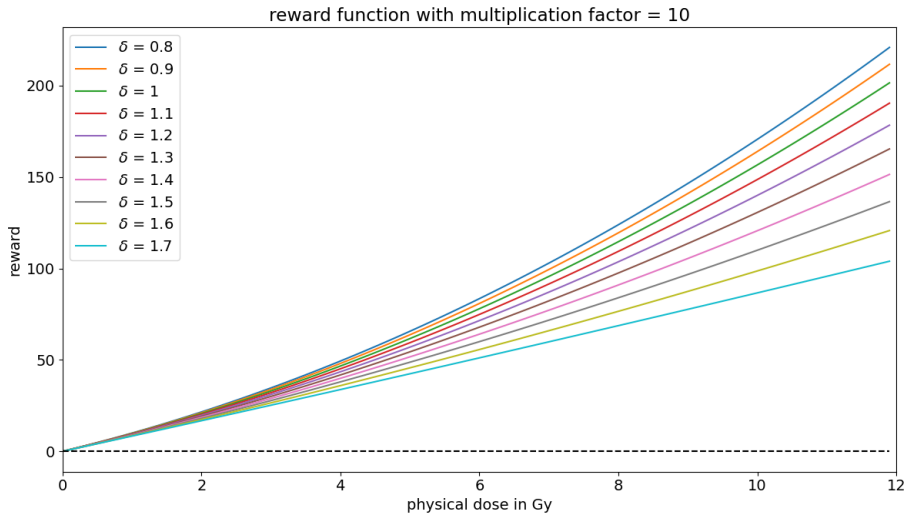
$$r_{diff} = f \cdot \left( d + \frac{d^2}{(\alpha/\beta)_T} \right) - \left( \delta d + \frac{\delta^2 d^2}{(\alpha/\beta)_N} \right) \quad (17)$$

As long the multiplication factor  $f$  is large enough,  $r_{diff}$  will always be positive and thus, the algorithm prefers to deliver more dose into the OAR instead of underdosing the tumor (except when the OAR dose surpasses  $B_{max}^N$ ). The resulting minimum value for  $f$ , so that the reward is positive for all delivered doses  $d$ , is given by:

$$f > \frac{\delta^2}{(\alpha/\beta)_N} \cdot (\alpha/\beta)_T \quad (18)$$



(a)



(b)

Figure 7: Reward properties based on sparing factor  $\delta$  and physical dose  $d$ . a) Is the reward function with a factor  $f = 1$ . b) Shows the behaviour at  $f = 10$

Following equation (18) a suitable multiplication factor  $f$  can be chosen. As the  $\alpha/\beta$  ratios are chosen to be constant in this work, the multiplication factor only depends on  $\delta$  which can

be regarded as the upper sparing factor limit, where the algorithm still prioritizes to maximize tumor dose instead of keeping the dose to be delivered to the OAR low. Figure 7 illustrates the properties of the reward function. If the multiplication factor  $f$  is not chosen high enough the reward becomes negative at high  $d$  values. We decided to use a multiplication factor of 10 which leads to an upper limit of sparing factors at 1.732. This guarantees that for all observed sparing factors improving tumor dose escalation will be prioritized.

---

**Algorithm 3** 3D algorithm

---

```

for t = F to 1 do                                ▷ Loop through all fractions starting from the last
  if t = F then                                       ▷ Initialization
     $B_{res}^N \leftarrow B_{max}^N - B^N$                                ▷ residual BED given by difference to reach  $B_{max}^N$ 
     $B_{res}^T \leftarrow B_{pre}^T - B^T$                                ▷ residual BED given by difference to reach  $B_{pre}^T$ 
     $d_1 \leftarrow d | \text{BED}_3(d, \delta) = B_{res}^N$ 
     $d_2 \leftarrow d | \text{BED}_{10}(d) = B_{res}^T$ 
     $d_{best} \leftarrow \min(d_1, d_2)$                                ▷ optimal dose given by the lower physical dose
     $\pi(t, \delta, B^N, B^T) \leftarrow d_{best}$ 
     $V(t, \delta, B^N, B^T) \leftarrow (B^T + \text{BED}_{10}(d_{best}) - B_{pre}^T) \cdot f - (B^N + \text{BED}_3(d_{best}, \delta))$ 
  else
     $B_{dose}^N \leftarrow \text{BED}_3(d, \delta)$                                ▷  $\text{BED}_3$  given by a physical dose  $d$ 
     $B_{dose}^T \leftarrow \text{BED}_{10}(d)$                                ▷  $\text{BED}_{10}$  given by a physical dose  $d$ 
     $V(t, \delta, B^N, B^T) \leftarrow \max_d [\sum_{\delta} P(\delta) \cdot V(t+1, \delta, B^N + B_{dose}^N, B^T + B_{dose}^T)]$ 
     $\pi(t, \delta, B^N, B^T) \leftarrow \text{argmax}_d [\sum_{\delta} P(\delta) \cdot V(t+1, \delta, B^N + B_{dose}^N, B^T + B_{dose}^T)]$ 
  return  $V, \pi$ 

```

---

The adjustment of the initialization with the final penalty  $v_{F+1}$  can be seen in code block 3, where the multiplication factor is depicted as  $f$ . Before computing the value in the last fraction, the optimal dose is computed by the action  $d_{best}$  which would precisely reach one of the two limits given by the OAR limit or the prescribed tumor dose and preventing the overdosage of both volumes. With the optimal dose decided for the last fraction, the value is assigned for the final state consisting of equation (16). In a next step, the interpolation between the discrete states is done similarly to the algorithms before with the only difference, that there are two BED values that are interpolated. The optimal doses for fractions 1 to 4 are more simple to calculate as no additional penalty or reward must be assigned. The optimal dose  $d$  is solely dependent on the future values  $V(t+1)$  and the probability distribution  $P(\delta)$ . To speed up the calculation of fraction one, the state space was limited to  $B^T = 0$  and  $B^N = 0$  as no dose should be accumulated in fraction one.

By tracking both OAR BED and tumor BED, the 3D algorithm has a clear advantage over the 2D-algorithms which can only pursue one objective. However, the extension of the state dependencies leads to more potential states that need to be calculated, consequently increasing computation time. A whole five fraction treatment plan can take several minutes to be calculated and is therefore limited in a potential online application.

### 3.4.3 Code extensions

Considering additional fractionation schemes and treatment limitations due to maximum dose constraints to the tumor in one fraction, two additional features have been added to all three algorithm types.

The first extension allows to compute treatment plans for an alternative amount of fractions. Since the MDP models consist of several subproblems, which are given by the different fractions, the increase or decrease in the amount of fractions does not change the architecture of the code and can easily be implemented. Nevertheless, the 3D algorithm is not well suited for a higher amount of fractions, as the computation time increases up to half an hour when evaluating fractionation schemes with for example 20 fractions. The 2D algorithms are still capable of computing treatment plans in a reasonable time when extending the number of fractions by keeping the computation time at a few minutes for treatment plans with 30 fractions.

A second extension was added to set maximum and minimum doses per fraction. In radiotherapy it might not be desirable to deliver little to no dose in a single fraction. Hence, a minimum dose can be chosen to assure that each fraction conforms with a standard dose. An upper limit guarantees that the optimal calculations do not surpass established fraction sizes.

The implementation of such constraints requires more adaptation of the original codes. In a first step, the action range is limited to the doses between maximum and minimum dose. Delivering a minimum dose in each fraction can lead to exceeding the OAR limit in the 3D algorithm and the 2D algorithm that maximizes tumor dose, which was limited by an infinite penalty or simply adapted by delivering the remaining difference to the objective. The algorithm would therefore deliver as much dose as possible in each fraction, as the penalty can not get more negative than  $-\infty$ . To bypass this problem, the overdosing penalty is not chosen as  $-\infty$  but, by a large negative number e.g.  $-10000$ . Additionally, a new penalty is introduced after the last fraction which is dependent on how much has been overdosed.

$$v_{F+1}(\delta, B) = \begin{cases} -10000 & (B > B_{max}^N) \\ 0 & (B \leq B_{max}^N) \end{cases}$$

$$r_{F+1} = -[(B^N - B_{max}^N)]_+ \quad (19)$$

If the algorithm is forced to overdose, the new penalty will lead to a minimization guaranteed by the overdosing penalty in equation (19) while overdosing is still prevented by all means as the fixed penalty is less desirable than any under-dosage of the tumor.

Another approach to modify the algorithm is to introduce heuristic techniques to potentially improve the code or lower the probability of getting an adaptive fractionation plan that is inferior than the reference treatment. Such a heuristic technique has been introduced with a risk factor  $R$ . Motivated by the idea to force the algorithm to stay closer to the reference treatment, the risk factor  $R$  introduces a new reward that is applied in each fraction  $t$ , except the last when  $t = F$ :

$$r_R = -|d_R - d_t| \cdot R \quad (20)$$

The dose  $d_R$  denotes the dose that would be delivered by the reference treatment in fraction  $t$ . With the additional risk factor penalty given by equation (20), a penalty is assigned to doses  $d_t$  that gets larger, the farther it deviates from the reference treatment. Consequently, the algorithm will additionally restrict the deviation from the reference dose to minimize the penalty. By altering the risk factor  $R$ , the restriction can be controlled. The larger  $R$  is chosen, the less the algorithm will deviate from the reference dose  $d_R$ . The risk factor can now be fixed at a

certain value or even be chosen to change dynamically based on the probability distribution of the sparing factors. An exploration of the effect of different risk factors showed, that risk factors above 1 always force the algorithm to follow the reference treatment. For this reason, a fixed risk factor of 0.1 was arbitrarily chosen for a larger analysis.

### 3.5 Probability updating

DP algorithms rely on a proper description of the environment to compute an optimal policy. In the application of adaptive fractionation, we do not have an exact description of the environment, the patient, as the probability distribution of the sparing factor  $P(\delta)$  is not precisely determined for each new patient. At the beginning of the treatment only little data is available to set up a probability distribution that describes a patient accurately. To improve the estimation of patient specific interfractional motion, one can rely on previous patient data to get a better initial probability distribution. As more information is gathered during the treatment, it is advantageous to integrate the patient specific information into the probability distribution.

For the sparing factor distribution a truncated normal distribution  $\delta \sim \mathcal{N}(\mu, \sigma^2)$  was chosen which has a lower limit at 0 and is defined by mean  $\mu$  and standard deviation  $\sigma$ . A first patient analysis displayed a large variation in sparing factor means between patients. Thus, the mean of the probability distribution is not dependent on a prior but calculated from the specific patient by updating the mean of the observed  $\delta$ . For any fraction  $t$  the mean of the sparing factor distribution is given by:

$$\mu_t = \frac{1}{t+1} \sum_{\tau=0}^t \delta_\tau \quad (21)$$

The variance of the distribution is often under- or overestimated if estimated from only two samples at the very beginning of the treatment. Therefore, the variance is computed via Bayesian inference after each new acquired sparing factor. Bayesian inference allows to calculate posterior probabilities based on a hypothesis  $H$  and evidence  $E$ .

$$P(H|E) = \frac{P(E|H) \cdot P(H)}{P(E)} \quad (22)$$

Here we want to compute the most probable variance  $\sigma^2$  that fits the observed data:

$$f(\sigma^2|\delta_0, \dots, \delta_n) = \frac{P(\delta_0, \dots, \delta_n|\mu, \sigma^2) \cdot f(\sigma^2; k, \theta)}{P(\delta_0, \dots, \delta_n)} \quad (23)$$

The hypothesis  $P(H)$  is the prior probability distribution of the variance  $\sigma^2$ , which includes information about the expected distribution and is described by  $f(\sigma^2; k, \theta)$ . On the left side of equation (23) is the posterior probability  $P(H|E)$ , the probability of  $H$  given  $E$ . When used to estimate parameters, the posterior probability describes how probable the hypothesis is, e.g. a model parameter, given the observed evidence. In this application, the probability of  $\sigma^2$  being an appropriate value to describe the observed sparing factors  $\delta_0, \dots, \delta_n$  (the evidence). Hence, to find the best fitting value, the posterior probability must be maximized. The probability  $P(E|H)$  is called the likelihood and indicates the compatibility of the evidence with the given hypothesis. To estimate the variance  $\sigma^2$ , we have to apply the probability of observing the sparing factors  $\delta_0, \dots, \delta_n$  given a mean  $\mu$  and  $\sigma^2$  which is defined by the probability  $P(\delta_0, \dots, \delta_n|\mu, \sigma^2)$ .  $P(E)$  contains the information derived from previous patients and must be chosen accordingly to the parameters that need to be estimated. Here, we can depict this probability as  $P(\delta_0, \dots, \delta_n)$  which normalizes the probability and can be dropped out as we are only interested in the maximum value of equation (23).

As the sparing factors are assumed to follow a normal distribution with unknown variance, an inverse-gamma distribution is chosen as prior distribution to compute the variance [27]:

$$f(\sigma^2; k, \theta) = \frac{\theta^k}{\Gamma(k)} (1/\sigma^2)^{k+1} \exp\left(\frac{-\theta}{\sigma^2}\right) \quad (24)$$

One can get rid of the scaling factor  $\frac{\theta^k}{\Gamma(k)}$  and estimate the hyper-parameters  $k$  and  $\theta$  with a maximum likelihood estimator by using patient data from the same population:

$$f(\sigma^2; k, \theta) \propto (1/\sigma^2)^{k+1} \exp\left(\frac{-\theta}{\sigma^2}\right) \quad (25)$$

Once the hyperparameters are known, the posterior probability  $P(H|E)$  is computed by multiplying the prior distribution (25) with the likelihood  $P(\delta_0, \dots, \delta_n | \mu, \sigma^2)$  which is defined by the normal distribution  $\mathcal{N}(\mu, \sigma^2)$ :

$$(\delta_0, \dots, \delta_n | \mu, \sigma^2) = \prod_{i=0}^n \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{\delta_i - \mu}{\sigma}\right)^2\right) \quad (26)$$

The resulting posterior probability is then:

$$f(\sigma^2 | \delta_0, \dots, \delta_n) \propto \frac{(\sigma^2)^{-(k+1)}}{\sigma^n} \exp\left(-\frac{\theta}{\sigma^2}\right) \exp\left(-\frac{1}{2\sigma^2} \sum_{i=0}^n (\delta_i - \mu)^2\right) \quad (27)$$

The parameters  $\delta_i$  denote the measured sparing factor at fraction  $i$ , where  $\delta_0$  is the sparing factor observed at the planning session.

The variance for each fraction can then be computed by maximizing the likelihood of equation (25):

$$\sigma_t^2 = \operatorname{argmax}_{\sigma^2} [f(\sigma^2 | \delta_0, \dots, \delta_t)] \quad (28)$$

With equations (21) and (28) the Probability distribution  $P(\delta_t)$  for each day can be updated.

### 3.6 Quantification of benefit

To quantify potential treatment improvement the treatment plans produced by adaptive fractionation are compared to three different treatments:

1. A reference treatment in which 6Gy physical dose (18Gy BED<sub>3</sub>) is delivered to the OAR in each fraction. Hence, the reference treatment delivers exactly the upper limit of 90 Gy BED<sub>3</sub> to the OARs. The reference treatment is not equal to the treatment plan that was delivered at the time of treatment of the discussed patients.
2. We calculate an upper bound for the benefit of adaptive fractionation. To do so, we consider the hypothetical situation that all sparing factors  $\delta_t$  are known before treatment. In such a case, the optimal dose for each fraction to maximize tumor BED is calculated by solving the following optimization problem:

$$\begin{aligned} & \underset{d}{\text{maximize}} && \sum_{t=1}^F \left( d_t + \frac{d_t^2}{\alpha/\beta_T} \right) \\ & \text{subject to:} && \sum_{t=1}^F \left( \delta_t d_t + \frac{\delta_t^2 d_t^2}{\alpha/\beta_N} \right) \leq B_N^{\max} \end{aligned} \quad (29)$$



The resulting treatment would optimally exploit the variation in  $\delta$  and can thus be used to benchmark the benefit of adaptive fractionation. However, it represents an unachievable upper bound for any realistic approach to adaptive fractionation where future sparing factors are unknown.

3. The clinically delivered plan at the time of treatment.

In the clinic, the delivered doses in each fraction are constrained by the maximum OAR BED but this dose is not a minimum dose to be delivered in each fraction. Hence, the executed plans deliver less than 6Gy to the OAR in some fractions. The reference plan was chosen to compare, what tumor doses can be reached, when fully exploiting the OAR constraint.

### 3.7 Patients and treatment plans

In the scope of this work we consider patients with abdominal tumors in proximity to either bowel, stomach or duodenum who received 5-fraction SBRT treatments at the MR-Linac system. The selection was motivated by the potential interfractional motion of abdominal tumors. In all cases tumor coverage was compromised due to the dose received by the dose limiting OAR. Hence, an adaptive fractionation treatment that escalates tumor dose may be of interest. The prescription doses and OAR constraints were not exactly equal throughout all patients. As a consequence, a uniform dose limit to the OAR has been set at 90Gy BED<sub>3</sub> to better compare the influence of adaptive fractionation. In general, the prescription doses and OAR constraints were quite similar and only deviated in a few Gray per fraction.

All patients were treated according to institutional practice. In addition to a pre-treatment planning MR-scan daily MR scans were acquired for online radiotherapy. Tumors and OARs were recontoured and daily adaptive treatment plans were created. Thus, the dose distributions were reoptimized in each fraction to adapt to inter-fractional changes without altering the prescription dose. For the purpose of this project, dose-volume histograms (DVH) of 16 patients were exported from the treatment planning system. DVHs were exported for GTV, PTV and the relevant OARs for six treatment plans, corresponding to the five delivered plans and the initial plan based on the planning MR. Based on this data the developed algorithms have been tested and analyzed.

## 4 Results and evaluations

### 4.1 Evaluated patients

Figure 8 shows the PTV sparing factor distribution for each patient and the respective dose-limiting OAR. Patients 1-6 were used to estimate the hyperparameters of the inverse-gamma distribution. Nonetheless, the patients were analyzed with the algorithms considering that there is a bias introduced by the conjugate prior for these six patients. Some patients show substantial variation with standard deviations of up to  $\sigma = 0.190$  (patient 13), whereas other patients show little variation of the sparing factor between fractions where the lowest standard deviation was observed in patient 2 with  $\sigma = 0.011$ . The respective csv-files with the DVHs for all patients can be accessed from a link in section 10.1. Additionally, the DVHs of patients 7,8 and 13 are provided in section 10.2 as well as the numerical values of the sparing factors for all patients.

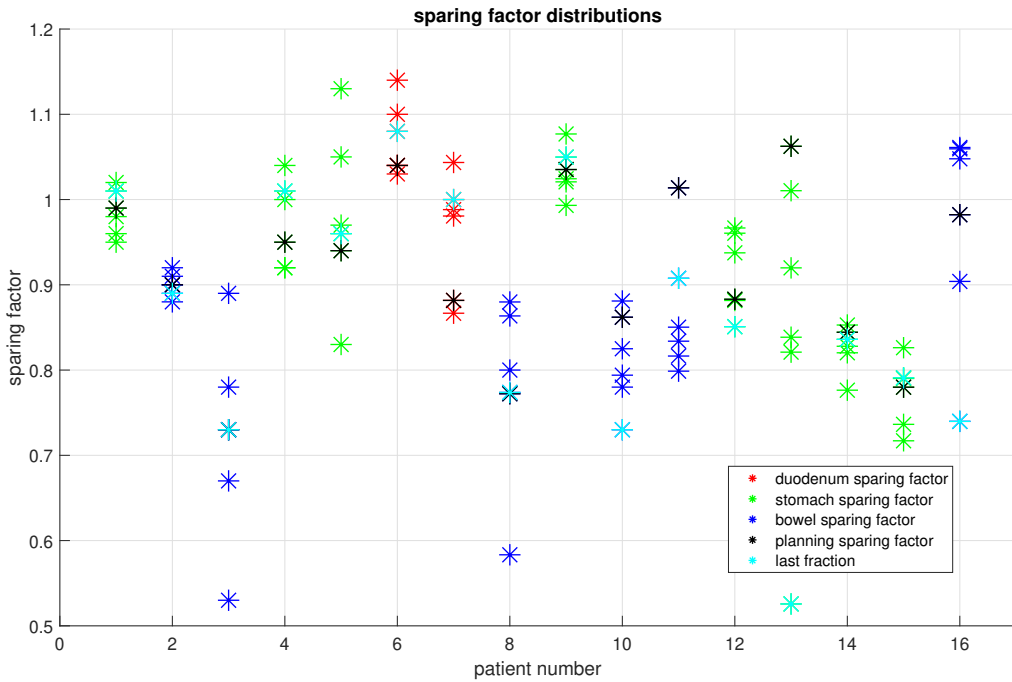


Figure 8: Scatter plot of all acquired sparing factors. The planning session and the last fraction have been highlighted.

In figure 8 the variation of the patient specific mean  $\mu$  of the sparing factors is visualized. Not just the standard deviation differs a lot in between patients, but also the patient specific mean. With the first six patients the hyperparameters of the inverse-gamma distribution were fit. The resulting distribution is shown in figure 9.

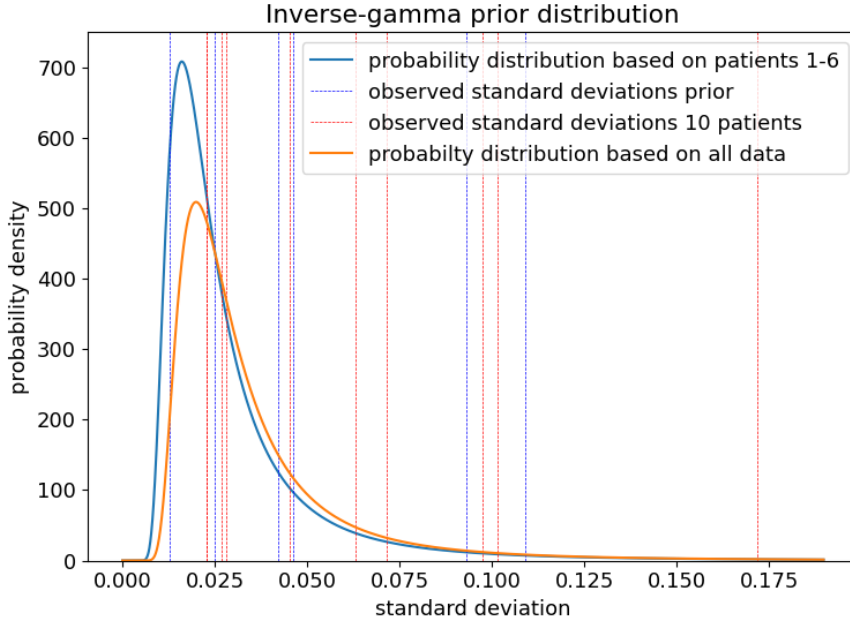


Figure 9: The inverse-gamma distribution used for adaptive fractionation is depicted in blue with the respective standard deviations of patients 1-6 marked with red lines. The inverse-gamma distribution based on all acquired patients is shown in the orange curve.

As an addition, all observed standard deviations were marked in figure 8. The distribution of standard deviations of the prior patients overlaps well with the distribution of the analyzed patients. One clear outlier can be observed which is the standard deviation of patient 13, which has a larger spread due to an extraordinary low sparing factor.

## 4.2 Illustration of adaptive fractionation for an example

To illustrate adaptive fractionation based on the DP algorithms, patient 7 is discussed thoroughly. Due to the high sparing factors and the compromise on the tumor dose for all fractions, the 2D algorithm that maximizes tumor dose subject to the maximum OAR dose is chosen for the analysis.

The sparing factors observed are  $\delta_0 = 0.88$  for the planning MR and  $[0.99, 0.87, 0.98, 1.04, 1]$  for the 5 fractions. In the first fraction, two sparing factors are known,  $\delta_0$  and  $\delta_1$ . These two sparing factors lead to an initial probability distribution  $P(\delta; \mu_1, \sigma_1)$  shown in figure 10a, with a mean of  $\mu_1 = 0.94$  and  $\sigma_1 = 0.036$ . Based on this probability distribution the values and optimal policies for all states can be computed by applying the Bellman equations (9),(10). The nature of the DP algorithm provides a solution for all fractions since the algorithm has to be initialized in the last fraction and then derives all states in a backwards sweep. In the fifth fraction the dose is defined by the difference between the maximum OAR BED and the accumulated dose of the fourth fraction, thus only fraction one to four are visualized. Starting from fraction two, the state space becomes two dimensional, since the accumulated OAR BED is non zero. Figures 10b,c,d depict the optimal doses based on the observed sparing factor and accumulated OAR BED. With the set of policies acquired from the initial sparing factor distribution a whole treatment plan can be computed (Figure 10a,b,c,d).

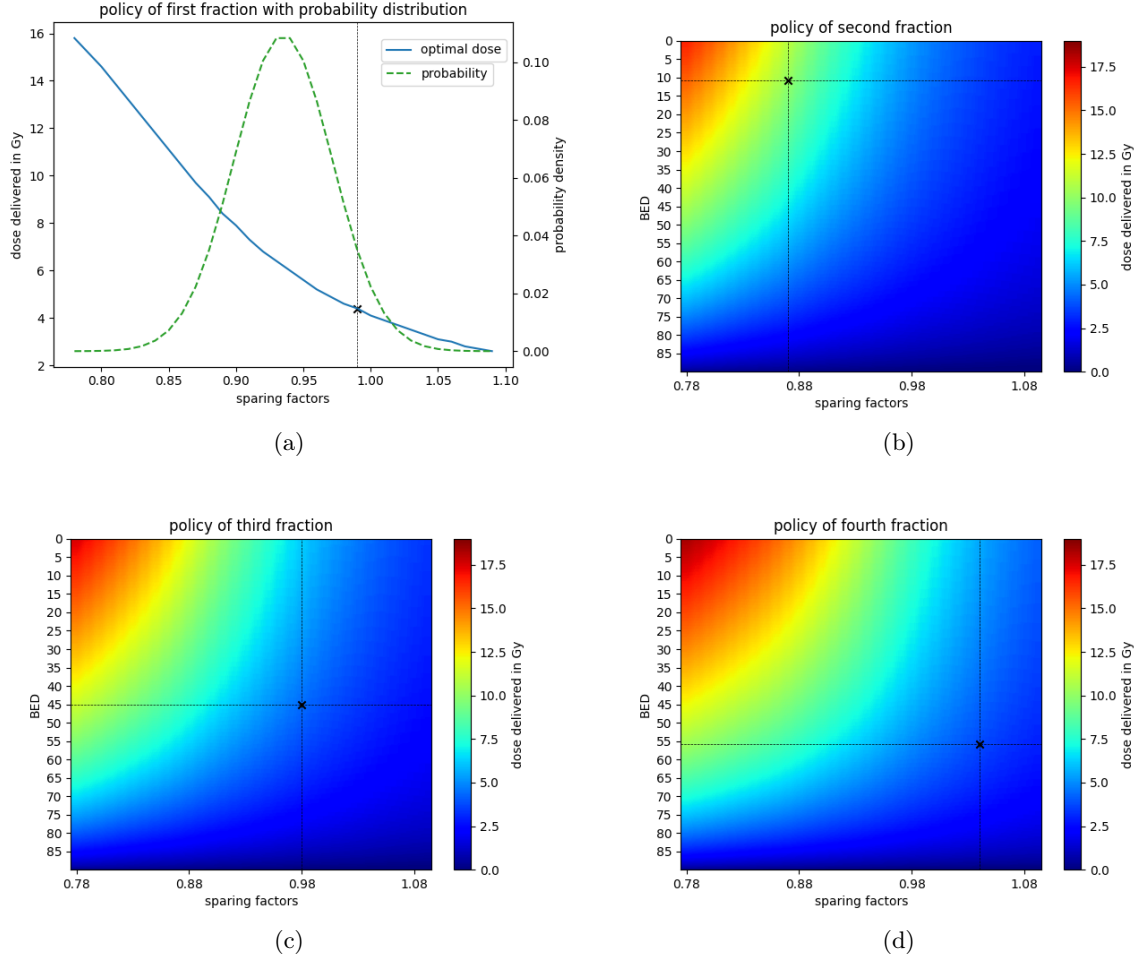


Figure 10: Initial policies of fractions one to four with the respective optimal policy, given that the optimal policy of the former fraction has been delivered

Since  $\delta_1$  is relatively high compared to the sparing factors to be expected given  $P(\delta; \mu_1, \sigma_1)$  a relatively low dose of  $d_1 = 4.4\text{Gy}$  is delivered to the tumor. The sparing factor in fraction two  $\delta_2 = 0.87$  is significantly lower. Hence, the optimal dose is now given by  $10.7\text{Gy}$  as depicted in Figure 10b. The third and fourth fractions are evaluated similar to the second fraction where the OAR BED is always given by the accumulated OAR BED so far.

In figure 10b,c,d it is clearly depicted that low sparing factors are favorable and lead to larger dose actions. Furthermore, low accumulated BEDs in the dose-limiting OAR are also preferred as that would mean that larger doses would not surpass the limit, nor would they prevent us from delivering doses in future treatments as the limit has not been exhausted. Therefore, a larger dose is suggested for lower accumulated OAR BEDs. Even though low accumulated BED states are favourable, to maximize the final accumulated tumor BED, the trade-off between landing in a low future OAR BED and delivering a large tumor dose must be optimized. To underline the impact of the observed sparing factor, the value function for the first fraction is illustrated in figure 11. Based on the sparing factor distribution set up by the observed sparing factors  $\delta_0$  and  $\delta_1$ , the expected tumor BED is computed to be  $52.4\text{ Gy BED}_{10}$ .

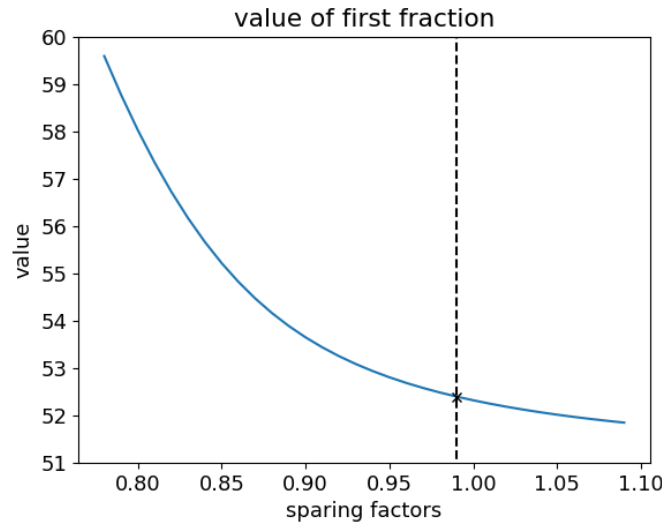


Figure 11: Values of the first fraction depending on the sparing factor. The sparing factor  $\delta_1$  is highlighted at the respective value of 52.4.

The dependence of the value on the accumulated OAR BED is illustrated in figure 12. The more BED has been accumulated, the less biological effective dose is expected to be delivered to the tumor since the bound of the OAR is reached earlier. Therefore, the states with low accumulated OAR BED are preferred.

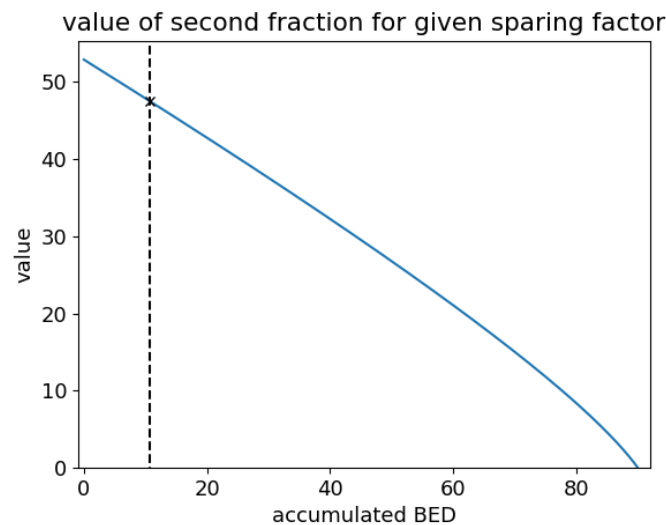


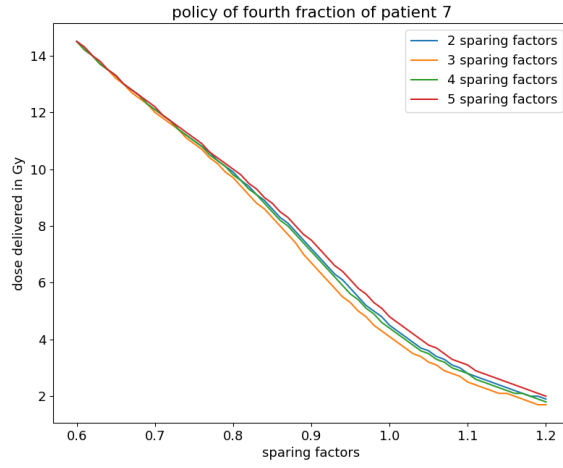
Figure 12: Values of the second fraction depending on the accumulated OAR BED with  $\delta = 0.87$ . The accumulated OAR BED resulting from fraction one is highlighted at the respective value of 47.5.

Integrating the acquired information about patient specific organ motion, the sparing factor probability distribution  $P(\delta)$  can be updated in each fraction. In fraction two  $P(\delta; \mu_2, \sigma_2)$  is given with  $\mu_2 = 0.91$  and  $\sigma_2 = 0.039$ . The new distribution has an impact on the computation of the optimal policy, thus all optimal doses for fraction two are recomputed. In this case the optimal dose for fraction two is given by 9.0Gy. The updating of the probability distribution is supposed to improve the description of the environment, the patient. Thus, the optimal plan is recomputed in each fraction and then compared to the reference treatments (Table 1).

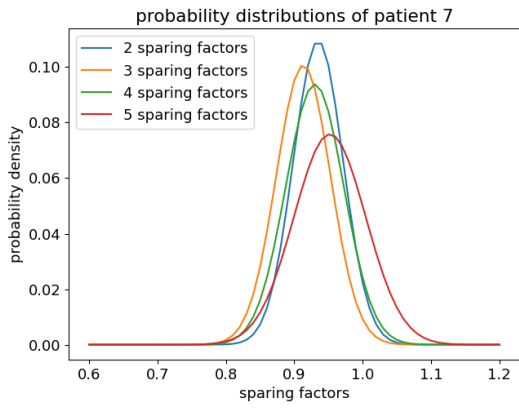
fraction	sparing factor	upper bound	adaptive fractionation	reference plan
first fraction	0.99	4.4	4.4	6
second fraction	0.87	12.6	9	6.9
third fraction	0.98	4.7	4.6	6.1
fourth fraction	1.04	3.3	4.3	5.8
fifth fraction	1.00	4.1	7.9	6
total BED tumor	-	51.7	50.4	50

Table 1: Dose delivered to the tumor based on different plans. The upper bound denotes the maximum tumor BED dose that could be achieved. The results for adaptive fractionation in this table are computed with probability distribution updating. Thus, the doses are not identical to the ones depicted in figure 10

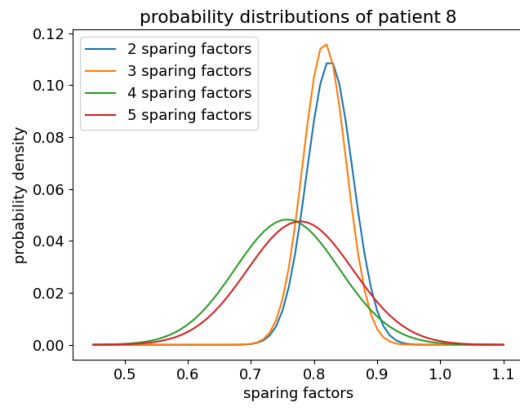
The modification of the optimal policy of fraction four throughout the treatment, when more information about the patient specific motion is integrated, is shown in figure 13a. Progressively adding information lead to optimal policies that differed up to 1Gy in fraction four. Figures 13b) and 13c) visualize the adjustment of normal distributions for patient 7, where all sparing factors are relatively close to each other and for patient 8, where an excessively low sparing factor was observed in fraction 3 resulting in a more substantial adaptation of the probability distribution. The final probability distribution of patient 7 has a slightly larger mean than the initial distribution which explains why the total BED delivered to the tumor is lower than expected in fraction 1 (value of 52.4). In fraction 3-5 the sparing factors were all above the mean of the distribution in fraction 1 resulting in a worse result than anticipated.



(a)



(b)



(c)

Figure 13: optimal policy and probability distribution progression.

### 4.3 Evaluation of patients

The 2D algorithm for tumor dose maximization has been run on all 16 patients and was then compared to the reference treatments bearing the following results:

patient number	reference plan BED	upper bound	adaptive fractionation	difference
1	49.1	49.4	49.2	0.1
2	55.6	55.7	55.6	0.1
3	81.0	108.7	98.7	17.7
4	49.7	50.6	50.2	0.5
5	49.7	53.9	51.7	2.0
6	43.4	43.8	43.5	0.1
7	50	51.7	50.4	0.4
8	71.3	93.1	84.4	13.1
9	46	46.3	46.2	0.2
10	67.3	68.8	67.3	0
11	61.2	62	61	-0.2
12	54.1	55.3	54.9	0.8
13	69.3	108.4	62.9	-6.4
14	63	63.6	63.3	0.3
15	69.1	70.9	70	0.9
16	52.5	63	51.3	-1.2

Table 2: comparison of treatment plans in Gy BED. The difference column is calculated as adaptive fractionation BED - reference plan BED. All delivered doses and sparing factors are listed in table 10

The delivered doses in each fraction for each patient are illustrated in figure 14 and all delivered doses and sparing factors are listed in table 10. Seven of the ten patients that have not been included in the prior calculation, have a positive difference while two of those show extreme results caused by outliers. No negative plans were computed for the six patients included in the prior estimation whereas patient 3 has an extreme result which is attributed by an extremely low sparing factor as well. The upper limit of adaptive fractionation for patients 1,2,6 and 9 was only 0.3-0.4Gy BED<sub>10</sub> above the reference plan. Consequently, adaptive fractionation did not have much room for improvement and the respective optimized plans are only slightly superior. In general, the difference between the adaptive fractionation and the reference plan is only around 1Gy BED<sub>10</sub> in the extracted patients. However, Patients 3, 8 and 13 show an exceptionally large deviation from the reference plan, which is owed to the larger variation in its sparing factors. For patient 8, a sparing factor of 0.58 is observed in the third fraction which is exploited by the delivery of 19.6Gy. This dose corresponds to most of the residual BED<sub>3</sub> that is allowed in the OAR. Thus, the algorithm expects worse sparing factors in future fractions, where delivering a large dose is not advantageous. A similar behaviour can be observed for patient 3, where a sparing factor of 0.53 shows up in the third fraction resulting in a physical dose to the PTV<sub>95</sub> of 20.7Gy. For both cases, the following sparing factors in fraction 4 and 5 are much larger than the sparing factor in fraction two. Therefore, delivering a large dose in fraction 3 was indeed a good decision, resulting in large improvements. Patient 7 has a substantial decrease in tumor BED when using adaptive fractionation. In this case the last sparing factor was exceptionally low. As the algorithm does not expect such a low sparing factor in the last fraction, the adaptive fractionation plan is inferior compared to the reference plan.



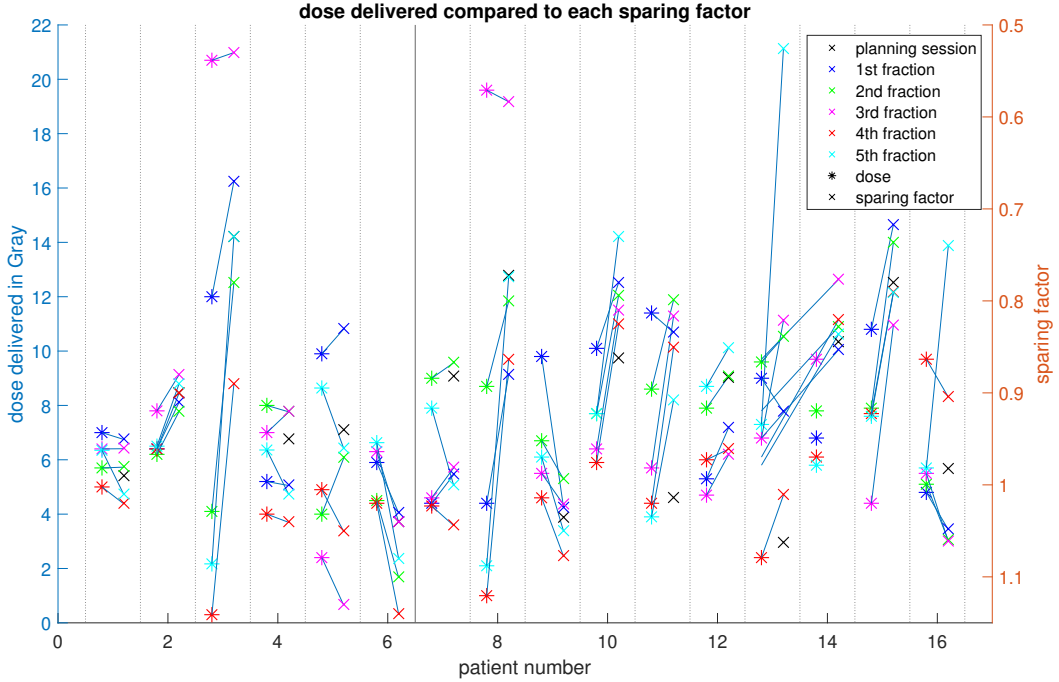


Figure 14: sparing factors and their respective doses for each fraction. Low sparing factors are often paired with high doses.

The clinically delivered treatment for patient 8 was based on a fixed prescription of 7Gy per fraction to the PTV and an OAR constraint of 6Gy per fraction. An illustration of the DVHs is given in figure 33. As the clinical treatment did not surpass the 7Gy per fraction to the PTV, the lower sparing factor in fraction three translates into a lower OAR dose instead of a larger PTV dose. In the clinical treatment, the OAR accumulated 76.4Gy BED<sub>3</sub> and stayed below the limit of 90Gy. As a consequence, only 59.2Gy BED<sub>10</sub> was delivered to the PTV. Hence, the prescribed dose of 72Gy could not be delivered to the PTV.

Similarly, the treatments for patient 13 (figure 34) can be compared to the clinically delivered treatment. The accumulated tumor BED was given by 58.6Gy BED<sub>10</sub>, 4.3Gy less than adaptive fractionation. Just as in patient 7, the inferiority is explained by the superior OAR sparing. At the time of treatment, only 79.9Gy BED<sub>3</sub> was accumulated in the OAR.

For all patients, adaptive fractionation always achieved a larger tumor dose, owing to the fact, that the clinically applied plans do not aim on the maximum OAR dose. A comparison with the clinically delivered plans therefore does not offer a good measurement for treatment plan improvement.

Optimal treatment plans, as they have been calculated in this subsection, can also be computed based on the GTV sparing factors. The resulting doses can then be applied to the PTV. Furthermore, the optimized treatment plans based on PTV sparing factors that have been described in table 2 can be applied to the GTV. Given that the reference plan is calculated with reference to the OAR the optimized treatment plans are compared to the same reference treatments as in table 10.

optimization type	GTV optimized	
Patient number	GTV difference to reference plan	PTV difference to reference plan
1	0.5	0.0
2	0.0	0.0
3	23.3	11.9
4	4.9	0.2
5	2.0	2.2
6	1.6	-0.1
7	-0.2	0.4
8	17.3	12.9
9	1.6	0.1
10	-0.1	0.0
11	-0.9	-1.9
12	1.6	1.2
13	8.7	3.9
14	0.4	0.0
15	2.1	0.7
16	2.4	-0.8
Mean	4.1	1.9
Mean unbiased	3.3	1.7

Table 3: Plan differences based on GTV optimization. The optimal doses have been computed based on the GTV sparing factors and then the difference to the reference plan has been calculated.

Optimization type	PTV optimized	
Patient number	PTV difference to reference plan	GTV difference to reference plan
1	0.1	0.4
2	0.1	0
3	17.7	32.9
4	0.5	2.5
5	2.0	1.5
6	0.1	-0.7
7	0.4	0.3
8	13.1	16.7
9	0.2	0.3
10	0	0.5
11	-0.2	4.9
12	0.8	1.0
13	-6.4	-8.2
14	0.3	-0.9
15	0.9	1.9
16	-1.2	-0.2
Mean	1.8	3.3
Mean unbiased	0.8	1.8

Table 4: Plan differences based on PTV optimization. The optimal doses have been computed based on the PTV sparing factors and then the difference to the reference plan has been calculated.

Comparing the results in tables 3 and 4, the mean of the differences is surprisingly larger for both PTV and GTV dose, based on the GTV sparing factors. Patient 13 stands out in table 3 with a benefit from adaptive fractionation. Based on the PTV sparing factors, an inferior plan was achieved with adaptive fractionation. Even though the lowest sparing factor is still in the last fraction in the GTV approach, the sparing factor in fraction 4 was higher than all previously observed sparing factors. Furthermore, the sparing factors of previous fractions did not have such a large variation as the PTV sparing factors. Thus, the GTV based adaptive fractionation was superior compared to the reference plan as a larger residual dose is available to be delivered in the last fraction.

To quantify the difference between the PTV and GTV sparing factors, their respective means and differences are listed in table 5.

Patient number	PTV mean	GTV mean	difference means	PTV std	GTV std
1	0.99	0.85	0.14	0.027	0.044
2	0.90	0.80	0.11	0.014	0.011
3	0.72	0.55	0.18	0.119	0.091
4	0.97	0.76	0.21	0.051	0.091
5	0.98	0.85	0.13	0.102	0.062
6	1.07	0.86	0.21	0.043	0.108
7	0.96	0.79	0.17	0.070	0.063
8	0.78	0.66	0.12	0.106	0.083
9	1.03	0.86	0.18	0.028	0.059
10	0.80	0.61	0.20	0.041	0.038
11	0.87	0.77	0.10	0.080	0.073
12	0.91	0.66	0.26	0.048	0.029
13	0.86	0.50	0.36	0.190	0.065
14	0.83	0.56	0.27	0.027	0.023
15	0.77	0.56	0.22	0.040	0.023
16	0.97	0.59	0.38	0.126	0.061

Table 5: Comparison of PTV and GTV sparing factors.

Over all patients the GTV sparing factors are lower than the PTV sparing factors but with no uniform scaling factor between the patients. In figure 15 the order of the sparing factors shows that if a PTV sparing factor in one fraction is larger than the PTV sparing factor in another fraction the same does not have to count for the GTV sparing factor. Hence, the optimality of a fraction is not unambiguous (e.g. fraction one and four of patient 8). Furthermore, the standard deviations of the two target volumes are sometimes larger in one set of sparing factors and sometimes in the other giving rise to potentially large differences in fractionation as it was the case for patient 13.

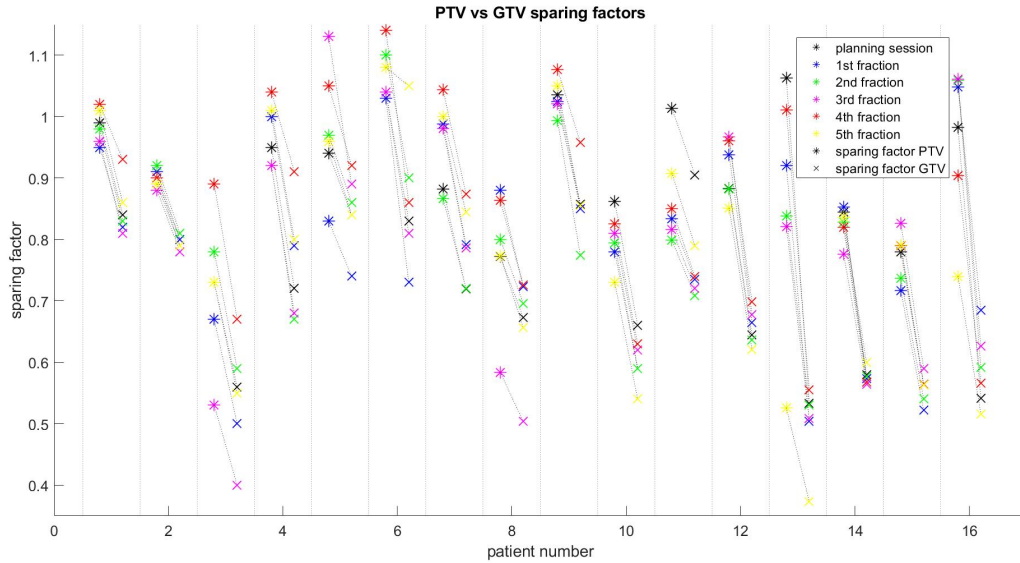


Figure 15: PTV and the corresponding GTV sparing factors of the unbiased patients.

#### 4.4 Permutation of extreme results

In table 2 the differences for patient 3, 8 and 13 are exceptionally large. To investigate the impact of the sparing factor order, when having exceptionally low sparing factors, the PTV based sparing factors have been permuted for patients 8 and 13 and all resulting adaptive fractionation treatment plans have been compared to the reference treatment plans with the new sparing factor order.

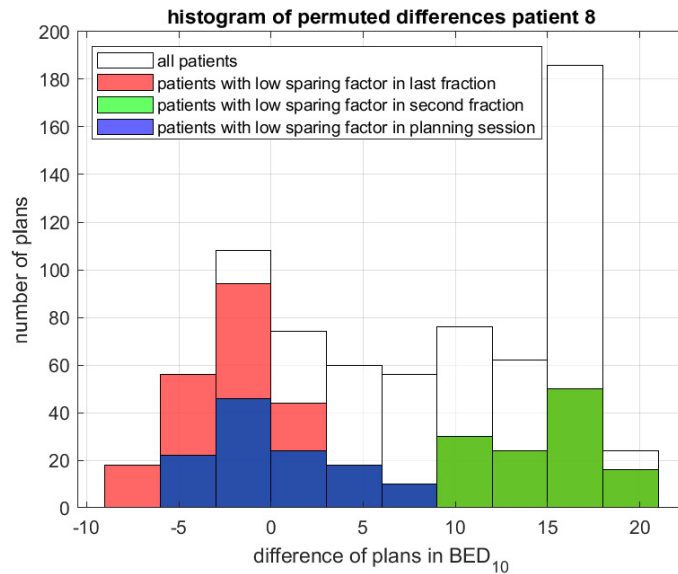


Figure 16: histogram of differences adaptive fractionation BED - reference plan. The blue bars show the results where the lowest sparing factor, 0.58, was at the last fraction and the red bars show the differences where the lowest sparing factor came up during the planning session. In green an optimal case is depicted where the lowest sparing factor was in the second fraction. The bars are stacked on each other to show the percentage of all patients in a certain bar. Results overall: 18.2% < 0

The permutation of the sparing factors of patients 8, [0.77, 0.88, 0.8, 0.58, 0.86], lead to the distribution shown in figure 16.

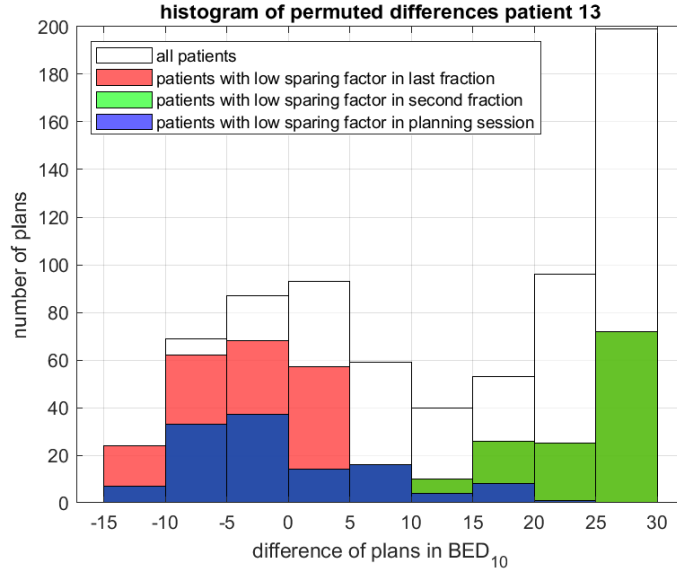


Figure 17: histogram of differences adaptive fractionation BED - reference plan. The blue bars show the results where the lowest sparing factor, 0.53, was at the last fraction and the red bars show the differences where the lowest sparing factor came up during the planning session. In green an optimal case is depicted where the lowest sparing factor was in the second fraction. The bars are stacked on each other to show the percentage of all patients in a certain bar. Results overall: 18.0% < 0

In figure 17 the sparing factors [1.06, 0.92, 0.84, 0.82, 1.01, 0.53] were permuted, where the special cases of having the lowest sparing factor at the planning session, in the second fraction or at the very end of the treatment have been highlighted. A clear trend of having worse results, when the exceptional sparing factor emerges in the planning session or in the very last fraction can be read from the diagrams. While an extreme sparing factor in fraction two always leads to positive results.

#### 4.5 Artificial data

Since the sample size of extracted patients is quite small, additional patient data has been produced by randomly drawing sparing factors from a predefined distribution to analyze the effect of adaptive fractionation. The artificial data has been chosen such that the distribution is similar to observed sparing factor distributions of the 16 patients. In a first step, two populations were considered where both populations have the same mean of  $\mu = 0.9$ . Two different standard deviations were used to simulate one cohort with large interfractional motion and one with more static patient geometry. Thus, one standard deviation was set at  $\sigma = 0.04$  and the other at  $\sigma = 0.07$ . Based on these distributions, 5000 patients have been drawn and a subset of 10 patients for each of these 5000 patients, which was used to define the hyper parameters of the prior distribution. The redrawing of the 10 prior patients assures that the results are not affected by one single randomly drawn prior. In figure 18 the differences between the adaptive fractionation and the reference plan for this setup are visualized. Additionally, 5000 randomly drawn patients have been analyzed, where the probability distribution was known and therefore, better results are to be expected.

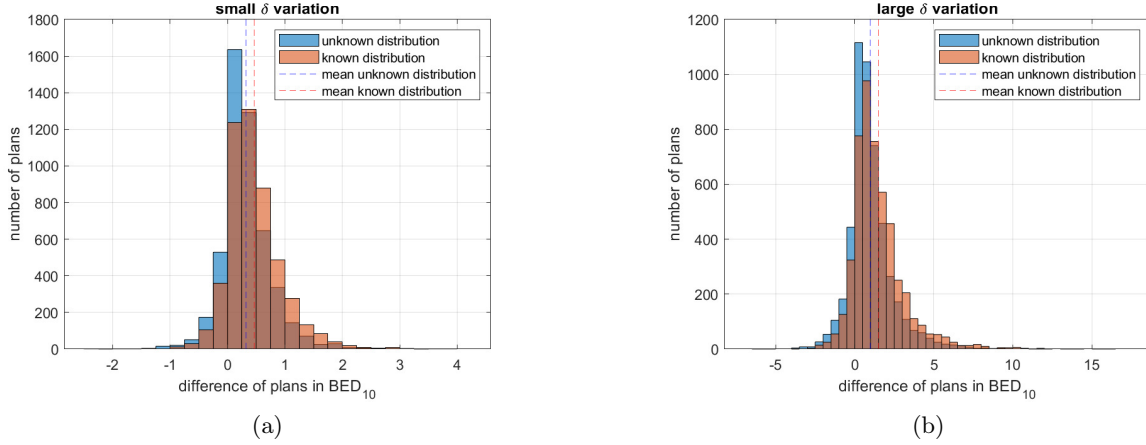


Figure 18: Differences of adaptive fractionation plan minus reference plan. a) unknown distribution: mean = 0.32, 15.9% < 0. known distribution: mean = 0.47, 10.1% < 0. b) unknown distribution: mean = 0.98, 16.7% < 0. known distribution: mean = 1.48, 11.0% < 0

To study the effect of a discrepancy between the prior population and the treated population, the prior data and the treated data have been drawn from different distributions. For this analysis, the same probability distribution as in the last example has been used. Since the mean of the prior data does not affect the prior distribution, only the standard deviation is chosen differently. An analysis, where the 10 prior patients are drawn with a standard deviation of  $\sigma = 0.04$  while the 5000 analysed patients have a standard deviation of  $\sigma = 0.07$  and one where the two standard deviations have been exchanged are visualized in figure 19.

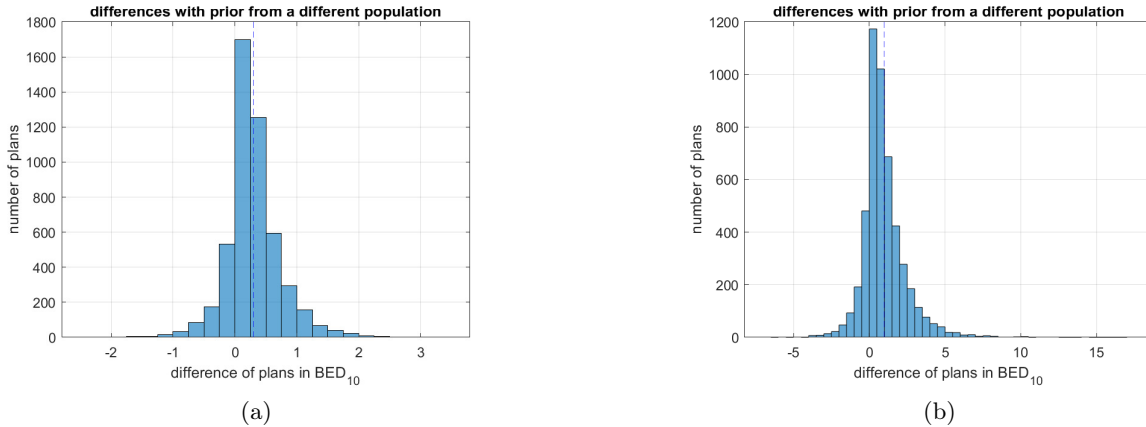


Figure 19: a) prior std = 0.07, treatment std = 0.04 17.0% < 0, mean = 0.30. b) prior std = 0.04, treatment std = 0.07, 17.3% < 0, mean = 0.96.

#### 4.6 Tumor dose objective and 3D algorithm

When a patient cannot be clearly assigned to a suboptimal anatomy geometry, where the tumor dose may be compromised, an optimization that considers OAR dose minimization is of interest. The analyzed cohort shared the trait of having compromised tumor dose thus, the 2D algorithm aiming on OAR dose minimization was not applied, but in the case of patient 13, an established tumor prescription dose  $B_{pres}^T$  of 72Gy BED<sub>10</sub> may be reached with adaptive fractionation. Thus,

the 2D OAR dose minimization algorithm can be applied. A comparison with the 3D algorithm which optimizes a treatment plan, following a prescribed dose of  $B_{pre}^T = 72\text{Gy BED}_{10}$  and a OAR constraint of  $B_{max}^N = 90\text{Gy BED}_3$ , is shown in table 6 for patient 8. The optimal doses based on the 2D algorithm for OAR dose minimization is given as well.

fraction	sparing factor	3D plan	2D plan	reference plan	upper bound
first fraction	0.88	4.4	4.8	6.8	0.4
second fraction	0.8	8.7	9.6	7.5	1
third fraction	0.58	17	16.3	10.3	21.6
fourth fraction	0.86	0.9	0.9	7	0.6
fifth fraction	0.77	2.1	1.9	7.8	1.3
total BED tumor		72	72	71.3	72
total BED OAR		77.7	79.9	90	68.3

Table 6: Dose delivered to the tumor in Gray. The 2D plan used the 2D OAR dose minimization algorithm applying the prescribed dose to the tumor. The optimal plan was computed with posterior knowledge, giving an upper bound.

Although the 3D and the 2D algorithm deliver the same total BED to the tumor, the accumulated BED in the OAR is not equal. The 3D algorithm is superior in this specific case. It can also be noted, that the reference plan can not reach the prescribed tumor dose subject to the OAR limit by applying 6Gy to the OAR in each fraction. The policies of the first fraction are given in figure 20. The 3D algorithm has a tendency to deliver less dose due to the different terminal reward in equation (16). Apparently, the additional penalty on an underdosed tumor and the upper bound on the OAR BED lead to the lower policy, as the probability distribution resulting from the sparing factors  $\delta_0 = 0.77$  and  $\delta_1 = 0.88$  expects lower sparing factors in future fractions which can be exploited better and thus improving the trade-off of the terminal reward.

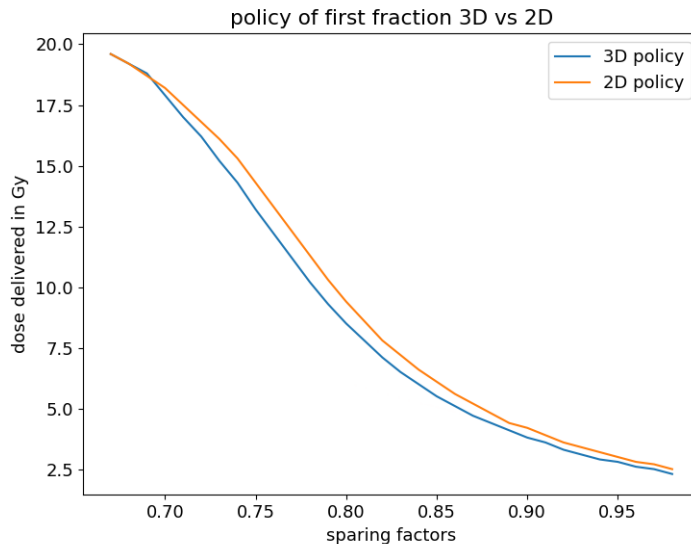


Figure 20: Comparison of first fraction policy for the 3D and 2D algorithm in patient 8.

In the third fraction, the 3D algorithm delivers a larger dose than the 2D algorithm which could imply that now the 3D algorithm tends to deliver larger doses in general. In figure 21 the policy of both algorithms are plotted against the accumulated tumor BED with the sparing factor fixed at  $\delta_3 = 0.58$ . The hypothesis that the 3D algorithm now applies higher doses can be proven

wrong, since the two algorithms are basically equivalent except for the first 10Gy BED. The reason for the higher dose of the 3D plan, is the lower accumulated tumor BED. As the 3D algorithm did not apply as much dose as the 2D algorithm in the former fractions, the optimal dose is higher for the 3D algorithm. Furthermore, the deviation at low BED values comes from the additional constraint of not surpassing the defined OAR BED. Thus, the 3D algorithm does not maximize the tumor BED as the 2D algorithm does for lower accumulated tumor BEDs.

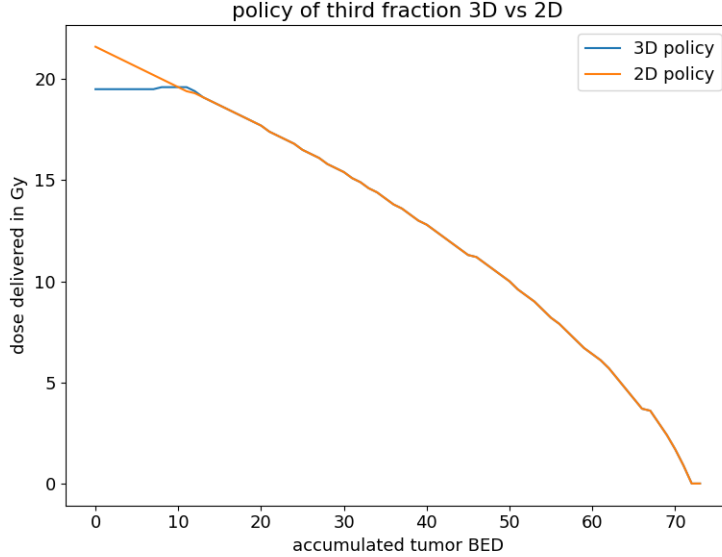


Figure 21: Comparison of first fraction policy for the 3D and 2D algorithm in patient 8, given that the sparing factor  $\delta_3 = 0.58$ .

Further, the 3D algorithm is applied to all other patients, where a theoretical prescribed dose of 72Gy BED<sub>10</sub> can not be reached. The results only deviate in patient 10 and 15 from the formerly computed plans with the 2D tumor dose maximization algorithm presented in the section 4.3

Patient 10			Patient 15		
sparing factor	3D plan	2D plan	sparing factors	3D plan	2D plan
0.78	10.2	10.1	0.72	11.9	10.7
0.794	7.7	7.7	0.74	7.5	7.7
0.81	6.3	6.4	0.83	4.2	4.1
0.825	6.2	5.9	0.79	6.6	7.1
0.73	7.3	7.7	0.79	7.9	8.7
total BED tumor	70.3	70		67.3	67.3

Table 7: Optimal dose differences for patient 10 and 15. The 2D plan corresponds to the tumor dose maximization algorithm.

Table 7 shows that for patient 10 a slightly larger tumor BED could be delivered, while for patient 15 the difference only becomes detectable with more significant digits, where the 2D plan would deliver more biological effective dose to the tumor. Additionally, the 3D algorithm starts with larger doses compared to the 2D algorithm. Here, the assurance of achieving the objective (reaching the prescribed tumor BED) is prioritized over minimizing the OAR BED from the very beginning. This behaviour highlights the impact of tracking the OAR and the tumor BED during the whole optimization.



## 4.7 Code extensions

### 4.7.1 Dose constraint extension

If there is no upper limit for a fraction dose the algorithm in theory could deliver unfeasible doses in a single fraction. In practice, proposed doses like 19.6Gy in fraction three of patient 8 (see table 10) would not necessarily be delivered due to lack of experience how tissue reacts to such large doses. Consequently, setting an upper limit for fraction sizes is of interest. A lower and upper limit that is motivated clinically would be given with  $d_{min} = 2\text{Gy}$  and  $d_{max} = 13.5$ . These limits were applied to the 2D algorithm to maximize tumor dose. In this case, patients 8 and 13 had altered plans where the total dose delivered to the tumor changed.

sparing factor	Patient 8		Patient 13		
	2D plan	2D constrained	sparing factor	2D plan	2D constrained
0.88	4.4	4.4	0.92	9	9
0.8	8.7	8.7	0.84	9.6	9.5
0.58	19.6	13.5	0.82	6.8	6.8
0.86	0.9	3.6	1.01	2.4	2.5
0.77	2.2	9.1	0.53	7.3	7.4
total BED tumor	93.1	76.7		62.9	63.1

Table 8: Differences between 2D optimization without constraint (2D plan) and with constraint. There was no difference in the constrained plans of the remaining unbiased patients.

The constraints lead to an decrease of tumor BED in patient 8, while increasing the tumor BED of patient 13 (see table 8). The remaining patients which were not included in the prior were not affected by the constraints as none of the previously optimized plans proposed to deliver a larger or lower dose than allowed.

As in section 4.5, two cohorts with 5000 patients have been generated where both populations have a mean  $\mu = 0.9$  and standard deviations  $\sigma = 0.04$  and  $\sigma = 0.07$ . In a first step, the resulting patients with a standard deviation of  $\sigma = 0.07$  have been evaluated with the constrained 2D algorithm where only a maximum dose  $d_{max} = 13.5\text{Gy}$  has been defined (figure 22). In a second step, the minimum dose  $d_{min} = 2\text{Gy}$  and the maximum dose constraint are applied on both cohorts. Constraining the maximum dose of the adaptive fractionation algorithm leads to a reduced mean benefit of 0.05Gy BED, but also to a decrease of inferior treatment plans. The reduction of mean benefit is mostly due to the decrease of extreme positive results. Similarly, limiting the actionspace by setting a minimum and a maximum dose results in slightly worse outcomes compared to the unconstrained plans for patients with small  $\delta$  variation, where a decrease of mean benefit of 0.01Gy BED has been observed, and also a decrease of superior treatment plans (figure 23). For the patients with a larger variation, the mean benefit now only decreased by 0.01 while the amount of superior plans increased. In figure 23 the difference between the adaptive fractionation plan and the reference plan are plotted, where the tails of the distribution are visibly less extreme for patients with large  $\delta$  variation when compared to the unconstrained plans in figure 18.

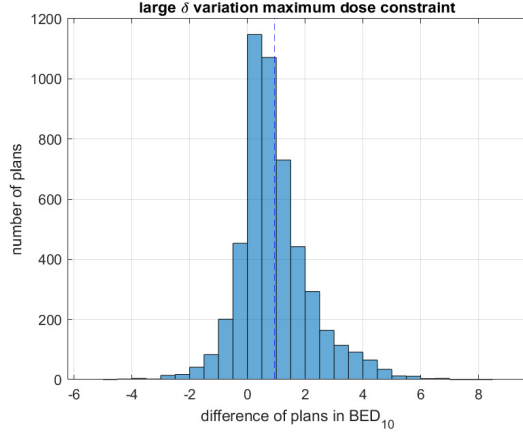


Figure 22:  $\delta$  distribution with 0.07 std and a maximum dose constraint of  $d_{max} = 13.5\text{Gy}$ . mean of differences = 0.93, 16.4% < 0.

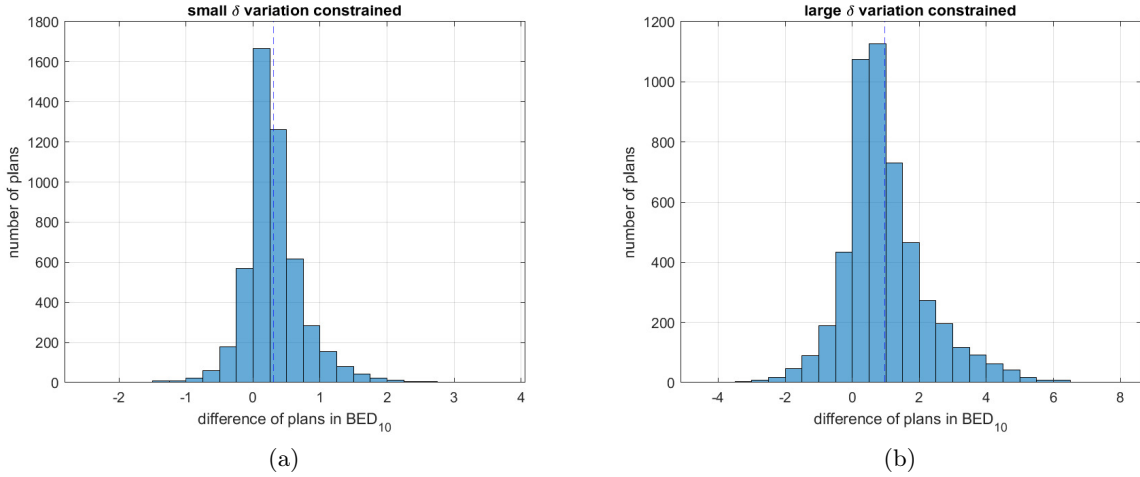


Figure 23: Differences of adaptive fractionation with minimum dose  $d_{min} = 2\text{Gy}$  and maximum dose of  $d_{max} = 13.5\text{Gy}$ . a)  $\delta$  distribution with 0.04 std. mean of differences = 0.31, 16.9% < 0 b)  $\delta$  distribution with 0.07 std. mean of differences = 0.97, 15.7% < 0.

#### 4.7.2 Risk factor extension

The impact of the risk factor is first analyzed on the extracted patients. In table 9 the differences between the plan optimization with the risk factor extension and the optimization without the extension are given. As expected, there are only minor differences. Patient 13, where an extremely negative result was obtained with adaptive fractionation, improved by 3Gy BED<sub>10</sub>, thus reducing the drawback of adaptive fractionation. By forcing the algorithm to stay closer to the reference fractionation, a larger residual dose is available in the last fraction which results in an improvement compared to the unconstrained adaptive fractionation.

Patient number	Difference
1	0.0
2	0.0
3	0.9
4	-0.4
5	-0.7
6	0.0
7	-0.1
8	0.2
9	-0.1
10	-0.1
11	0.6
12	-0.4
13	3.0
14	0.0
15	-0.5
16	0.1

Table 9: Difference of adaptive fractionation with a risk factor of 0.1 minus adaptive fractionation without risk reduction.

An extension to 5000 generated patients showed a decrease in mean benefit of the treatment plans. For patients with large variations, the extreme negative plans were limited similarly to the plans with a dose constraint (see figure 24). But additionally, some of the extreme plans with large benefit were conserved. Similarly, the spread of the differences was reduced for the small variation where the most extreme inferior plan was given with a difference of -0.6Gy BED. Also, there is a more notable decrease of treatment plans with large benefit.

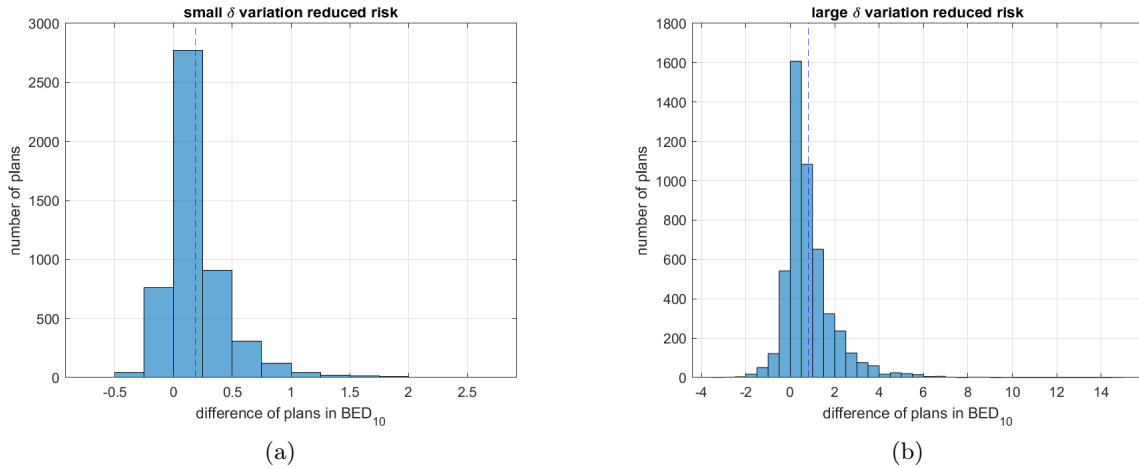


Figure 24: a)  $\delta$  distribution with 0.04 std. mean of differences = 0.19, 16.1%<0 b)  $\delta$  distribution with 0.07 std. mean of differences = 0.82, 14.8%<0

## 5 Graphical user interface

To grant easy access to the DP algorithm, three different interfaces have been developed. Two of them allow direct plan calculation where the 2D functions and 3D function have been split into different programs with similar architecture.

Figure 25: 3D interface with default values set when opening the program.

Figure 25 shows the general layout of the interface which is split into four main blocks:

- 1 Probability distribution menu to define the features of the probability distribution
- 2 Extension menu to set the number of fractions and dose constraints
- 3 Main input menu. In this block the patient specific information is provided
- 4 Output window where the resulting optimal doses are displayed

All of the input blocks have an information button joined, marked by a question mark to get further information how to fill in or use the respective feature. Additionally, both of the calculation interfaces come with default values set adapted to the patients used in this work.

## 5.1 Probability distribution menu

The first main block includes three different options to set the probability distribution to be used in the dynamic programming algorithm to compute the values and optimal policies (see equations (9) and (10)). The standard function is to use hyperparameters estimated from prior patients that are assumed to be from the same population as the analyzed patient. To give visual feedback on the impact of the hyperparameters, a third interface has been provided, which plots the resulting distribution of the conjugate prior thus allows a manual setup of the hyperparameters even without prior data.

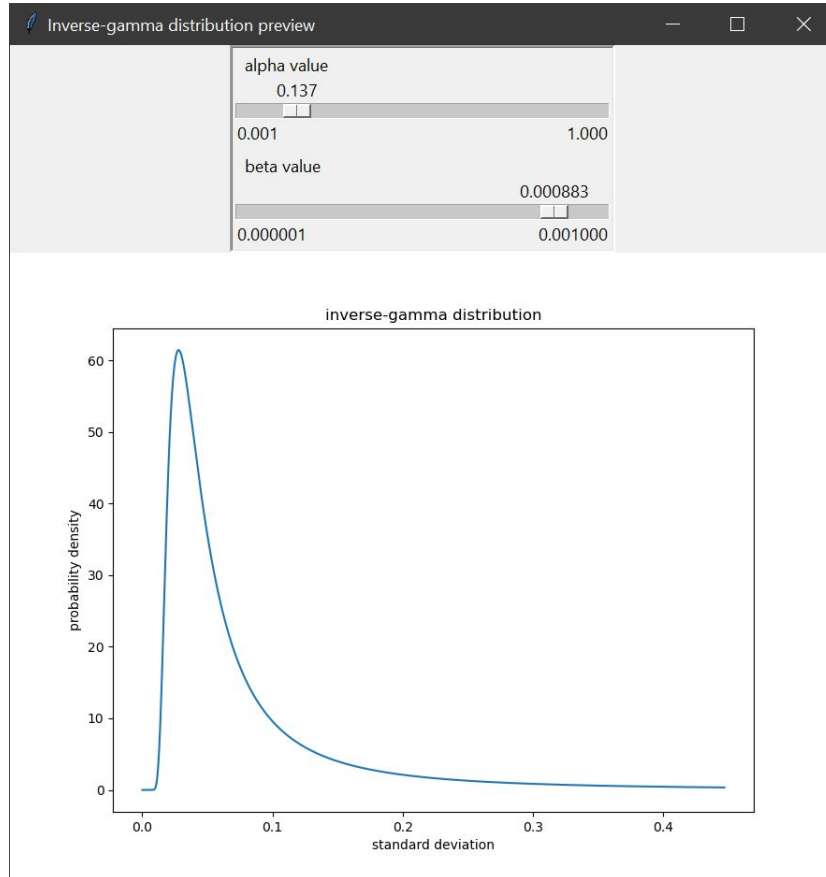


Figure 26: Inverse-gamma distribution interface

In the interface (figure 26), two sliders allow the manipulation of the conjugate prior plotted below, thus giving a direct visual feedback about the prior distribution to be used in the optimization.

A second option is to use prior data to directly estimate the hyperparameters. To do so, a csv file can be selected via the *Open a file* button that is enabled if the radio button *prior data* is selected (see figure 27). The data must be ordered in such that each row represents a new patient while the different sparing factors observed for a patient are separated in columns.

Figure 27: Probability distribution block with prior option selected. Changing the distribution setup via the radio buttons enables the respective options and disables the other input areas.

If the sparing factor distribution should not be updated, but the whole plan should be computed by a fixed probability distribution, the *define normal distribution* radio button can be selected. This option enables the setup of a normal distribution by setting the mean and standard deviation of the sparing factor distribution that is used for all fractions of a patient.

## 5.2 Extension menu

The default values of the extension menu are set to a five fraction treatment with no maximum or minimum dose in each fraction. Increasing the number of fractions leads to longer computation time, which are still in time frame of a few minutes even at 30 fractions for the 2D algorithm. The computation time for the 3D algorithm is heavily increased. Thus, the computation of 30 fraction treatment is not recommended with the 3D interface.

Setting up minimum and maximum doses to be delivered in each fraction lower computation time. Hence it is advised to set additional boundaries if the 3D interface is used with more than five fractions.

### 5.3 Main input menu

As the 2D and 3D programs do not use the same objectives, the main input menu differs between the two interfaces.

(a)

(b)

Figure 28: a) 2D interface main input menu. b)3D interface main input menu.

The five entry areas starting from the top are the same for both interfaces. A list of sparing factors must be provided to compute optimal doses. The separation of the sparing factors must be done by spaces. Any other separator leads to an error. alpha-beta ratios, OAR limits and prescribed tumor doses can be given as float numbers. For the 2D interface, the prescribed tumor dose is disabled in the default operation. The 2D algorithm maximizes tumor dose subject to an OAR limit, thus a prescribed tumor dose is not needed. A second 2D algorithm can be used, to minimize OAR BED and aiming on the prescribed tumor dose. To do so, the *Calculate optimal plan by minimizing OAR and aiming on prescribed tumor dose* checkbox must be selected. The checkbox disables the OAR limit and enables the prescribed tumor dose in the 2D interface and changes the optimization algorithm to be used.

If the optimal doses for a single fraction, instead of a whole plan, should be calculated, the algorithm requires the amount of dose that was delivered to the OAR and tumor. The 2D interface only needs the accumulated dose of the tracked tissue, i.e. the OAR BED in default mode and the tumor BED if OAR dose should be minimized. To enable the calculation of a single fraction, the *Calculate dose only for actual fraction* must be selected. The actual fraction is given by the amount of sparing factors. For example, if the third fraction dose should be calculated, all four known sparing factors are to be provided as depicted in figure 28a.

## 5.4 Output window

In the output window, all results are displayed. The display type depends on whether one fraction is calculated:

compute plan	
Optimal dose for fraction 3 = 4.1000000000000005	
delivered tumor BED = 5.7810000000000015	
delivered OAR BED = 10.32456533333334	

(a)

fraction number	sparing factor	physical dose delivered to PTV95	BED delivered to tumor	BED delivered to OAR
fraction 1	0.99	4.4	6.34	10.68
fraction 2	0.87	9.0	17.1	28.27
fraction 3	0.98	4.5	6.52	10.89
fraction 4	1.04	4.3	6.15	11.14
fraction 5	1.0	7.95	14.27	29.02
accumulated doses			50.38	90.0

(b)

Figure 29: a) Output for single fraction computation. b) Output for a whole five fraction treatment.

Since the 3D algorithm takes a few minutes to calculate a whole five fraction treatment, a progression bar has been implemented in the output window. The progression bar depicts the amount of fractions that have been calculated. Therefore, each progression of the bar takes less time as the farther into the treatment, the less future fractions have to be considered (see figure 30).

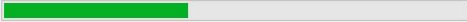
compute plan	
There are several default values set. Only the sparing factors have to be inserted.	
This program might take some minutes to calculate	
	

Figure 30: 3D interface progression bar.

The 2D interface takes only a few seconds to compute a complete five fraction plan and roughly three minutes for a 30 fraction plan with updated probabilities. Hence, no progression bar was implemented.



## 6 Discussion

### 6.1 Illustration of adaptive fractionation for an example patient

Table 1 shows that the 2D algorithm has the potential to approach the optimal treatment plan and therefore increasing the tumor dose during the treatment. Furthermore, the impact of probability updating depicted in figure 13 displays that the addition of information is valuable, as the decision making has an impact of up to one Gray in fraction 4 for patient 7, where the sparing factors have a standard deviation of 0.06. The optimal dose of fraction one is decreased by 1.7Gy even though a higher dose would be more optimal. Thus, the improved probability distribution does not necessarily mean that the decision making is shifted towards the most optimal plan by updating the probability distribution. Nevertheless, the adjustment of the probability distribution for patient 8 illustrates how big of an impact the probability updating can have to describe the potential motion of a patient.

### 6.2 Evaluation of patients

An analysis of all patients with the 2D algorithms is listed in table 2. Patients 3,8 and 13 stand out the most. The large differences to the reference plans are explained by their sparing factor distributions which is depicted in figure 13. The three patients have one exceptionally low sparing factor, where patient 3 and 8 have the exceptional sparing factor between fraction one and four whereas patient 13 has the low sparing factor in the fifth fraction. A low sparing factor between fraction one and four allows an optimal exploitation. If the extraordinarily low sparing factor shows up in the fifth fraction, the exploitation is not possible. The adaptive fractionation algorithm does not "expect" such a sparing factor based on the probability distribution so far in the planning session and the fractions one to four. Thus, the algorithm does not save a large amount of dose for the last fraction and ends up delivering less dose than the reference plan, that delivers the same dose to the OAR in each fraction. Hence, the results are worse.

Considering the less extreme patients, the DP approach seems to be roughly equal to the reference approach. The patients included in the prior calculation improve slightly when adaptive fractionation is applied while three of the ten unbiased patients have negative results. In general, the differences are distributed closely around zero if no exceptional sparing factors show up. A potential problem of the 2D algorithm that only tracks OAR BED is a possible overdosing of the tumor. patients 3 and 8 reach rather high tumor BEDs which could be problematic in a clinical setup. To handle such overdosing, the clinicians could limit the doses that are to be delivered during the application of the treatment or stop the treatment once a prescribed tumor dose has been delivered.

To investigate whether there is a difference in treatment plan quality if GTV based sparing factors are used for an optimal treatment calculation, the GTV based plans were compared to the PTV based plans in tables 3 and 4. Judging from the results, it seems to be advantageous to use the GTV to compute an optimal fraction dose as the average benefit of both target volumes, GTV and PTV, is larger in that case. As the sample size is quite low, a paired t-test was applied on the unbiased ten patients to check whether the plan differences are significant. The differences of the PTV plans have a p-value of 0.85 and a 95%CI of [-1.5,1.8]. The differences of the GTV plans have a p-value of 0.57 with a 95% CI of [-2.07,3.6]. Thus, both results are not significant and the CIs include zero. It follows, that these results could be coincidence and a larger sample size is required to get a significant result. A further analysis to check whether there is a regularity between the PTV and the GTV based sparing factors in table 5 did show the expected tendency that all GTV based sparing factors are lower than the PTV based. This is due to the higher prescription dose to the GTV and therefore originates in the definition of

the PTV as an expansion of the GTV. Given the observed data, there seems to be no simple transformation between the PTV and GTV sparing factors. Even the standard deviations of the sparing factor distributions did not correlate on a regular basis. A first assumption, that the standard deviations are roughly the same for PTV and GTV sparing factors was proven wrong. Since the delineation of the target volumes follows a complex process that can not be modeled easily with a probability distribution, no further analysis with artificial data was pursued on this topic.

### 6.3 Permutation of extreme results

The impact of extreme sparing factors and the fraction in which they come about was further studied by permuting the sparing factors of patient 8 and 13. Both outcomes have been visualized in figures 16 and 17. By highlighting the differences, the assumption, that extreme negative results occur when the exceptionally low sparing factor is in the last fraction was affirmed. Additionally, having the low sparing factor in the planning session may lead to bad results as well. By observing an extremely low sparing factor in the planning session, the algorithm underestimates the mean of the future sparing factor and proposes doses that are not well suited for the future sparing factors. Having the exceptional sparing factor in fractions 1 to 4 can be exploited as suggested before. Especially in the second fraction, the algorithm can still adapt to such a low sparing factor as a lot of dose has yet to be delivered where results with a difference up to 30Gy BED can arise.

### 6.4 Artificial data

The artificial data shows that a larger standard deviation in a population leads to more extreme results. A larger variation between the sparing factors gives rise to a potentially better plan, but also a larger risk of getting a significantly worse treatment plan. Furthermore, exact information about the probability distribution increases the benefit of adaptive fractionation by producing less negative plans, i.e. less plans that are inferior compared to the reference treatment, and increases the mean of the differences. In contrast to the 16 analyzed patients, the amount of negative plans is below 18% for both cases, where the prior patients are drawn from the same distribution as the analyzed cohort and where the prior patients have a different probability distribution than the analyzed data.

A closer look on the analysis where the prior population was drawn with a different probability distribution in figures 19a and b shows that the quality of the results only decreases marginally compared to the case where the prior was chosen to coincide with the analyzed patients. In the case, where the prior had a smaller standard deviation than the actual population, only a increase of 0.6% in negative cases has been observed as opposed to the the increase of 1.1% when the standard deviation is overestimated. This suggests, that underestimating the spread of the sparing factors has a lower impact on the results. Generally, the prior distribution does not have such a large impact on the adaptive fractionation results when dealing with standard deviations below 0.1.

### 6.5 Tumor dose objective and 3D algorithm

Table 6 lists a case of a patient where a prescribed tumor dose of 72Gy BED<sub>10</sub> can be reached. There is a slight discrepancy between the 2D plan that aims on the prescribed tumor dose and the 3D plan that optimizes tumor and OAR dose. The 3D plan achieves a better result in this case. This can not be generalized for all patients based on this result. Generally, similar results are to be expected if the prescribed tumor dose can be reached based on the same penalty that is assigned to the OAR dose. Nonetheless, from the listed examples, tracking tumor and

OAR BED is superior than just optimizing based on one single objective since borderline cases, where the achievement of the prescribed tumor dose is not certain, can be properly optimized. To further examine the 2D algorithm that minimizes OAR dose and compare it with the 3D algorithm, patients where the tumor dose did not have to be compromised need to be exported. A generation of artificial patients with such preconditions has not been conducted.

The application of the 3D algorithm on the patients, where a prescribed tumor dose of 72Gy BED<sub>10</sub> could not be met lead to similar results as the 2D algorithm in section 4.3. Since the rewards of the 3D algorithm are setup in a related manner, the results only differ in cases where there are rather low sparing factors and the prescribed tumor dose is nearly achieved as in patients 10 and 15.

## 6.6 Algorithm extensions

Since this work focuses on patients who undergo SBRT treatments at the MR-Linac, a thorough discussion of cases where a patient is treated in more fractions is omitted. In the clinic, patients with more fractions are generally not treated at the MR-Linac, as the high precision is often not required, thus the daily scans are not provided. In addition, the workload at the MR-Linac is high, thus extended fractionation schemes can not be considered. Evaluations of treatments with larger fraction numbers have been covered by Ramakrishnan et al.[22] but with artificial data.

A reasonable upper and lower limit for the dose per fraction was given with 2Gy and 13.5Gy. Such a limitation reduces the action space of the algorithm, altering the results of patient 8 and 13 (see table 8). If tumor dose escalation was enabled by exploiting a low sparing factor, as it is the case with patient 8, the escalation is limited by the upper bound. Furthermore, the lower bound forces a minimum dose, which increases the dose to be delivered in fraction 4, deviating further from the optimal treatment given in table 10. The resulting treatment plan delivers 16.3Gy BED<sub>10</sub> less to the PTV. In other cases, where the optimal sparing factor may show up in the last fraction, e.g. patient 13 the lower limit can lead to improved plans, compared to unlimited adaptive fractionation. The lower limit assures that a larger dose can be delivered in the last fraction, thus improving the treatment plan. By increasing the lower limit to 4Gy, the accumulated tumor BED increases to 65.6Gy BED<sub>10</sub> reducing the tumor BED decrease to -3.7Gy BED<sub>10</sub>.

An evaluation of 5000 generated patients with small and large standard deviation indicated a slight decrease of treatment plan quality compared to unconstrained adaptive fractionation. In the population with a standard deviation of 0.04, the mean benefit was reduced by 0.01, while 1% more patients had a negative treatment plan. The population with a standard deviation of 0.07 had a similar decrease in differences of 0.01, while the number of patients with an inferior plan decreased from 16.7% to 15.7%. Compared to just setting an upper dose limit, the application of a minimum and a maximum dose delivers better results as the mean benefit decreases by a smaller margin and the amount of inferior decreases to 15.7% instead of just 16.4%. A graphical comparison in figure 31 shows an decrease of spread of the differences for both populations, while the population with large standard deviation has mostly less extreme outliers. In summary, the proposed limitation of the action space has a negligible impact on the mean benefit of the patients, while patients with large  $\delta$  variation have a reduced probability for an extreme inferior or superior treatment plan.

Introducing a heuristic constraint to reduce the deviation from the standard plan by adding a penalty had only a minor impact on the extracted patients. 7 out of 16 patients had a decrease of benefit while the adapted treatments were still superior to the reference treatments. The benefit for patient 11 even increased such that the optimal plan based on the reduced risk algorithm became better than the reference treatment, which was not the case for the unconstrained plan.

Applying the algorithm on the artificially generated data shows that the reduced risk model does yield the desired effect of lowering the percentage of negative plans for larger  $\delta$  variation and also reduces the amount of extreme inferior plans but with the cost of also limiting the large benefit from adaptive fractionation. Nonetheless, more positive outliers can be observed for the larger  $\delta$  variation in 31 than when using the dose constraint.

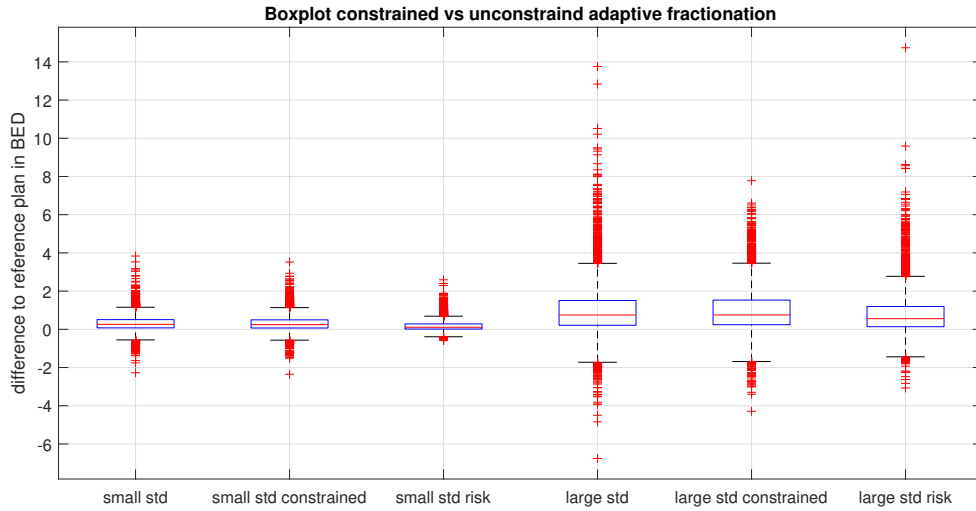


Figure 31: Boxplot summarizing the distributions from figures 19, 23 and 24.

## 7 Conclusion

Using a DP algorithm for adaptive fractionation when there is a tumor and organ movement involved during the treatment on average only provides a small increase in tumor dose in patients where tumor dose had to be compromised. By sampling a larger number of patients, a tendency toward a slight mean increase in treatment quality when using adaptive fractionation under the assumption that the prior distribution is not excessively different compared to the true population distribution can be shown. A larger impact on the treatment quality arises from extraordinarily low sparing factors. Having such exceptional sparing factors in the planning session or the last fraction significantly worsen the treatment plans when no constraints are applied to the algorithm. Whereas extremely low sparing factors in fraction one to four can be exploited and result in largely improved treatment plans. In general, patients with large variations in geometry may yield substantial benefits when adaptive fractionation is applied.

The application of the 3D algorithm can further improve the choice of doses to be delivered as an optimal prescription dose can be pursued without risking an overdosing of the OAR. Hence, the 3D algorithm is superior compared to the two 2D algorithms, as the algorithm delivers nearly identical results for patients where the prescribed dose is reached throughout the treatment and where the goal is to minimize the OAR BED such as for patients where the OAR BED is reached and tumor BED is to be maximized without having to decide on one of the two algorithms at the beginning of the treatment. A drawback of the 3D algorithm is the larger computation time, which could easily be reduced by using a parallel programming approach which is well fit for dynamic programming. Nonetheless, the optimal dose for a single fraction can be calculated in a few minutes for a five fraction treatment making the algorithm potentially applicable in the clinic.

Adding further constraints on the 2D algorithm to maximize tumor dose limits the benefit of adaptive fractionation. An upper limit of 13.5Gy restricts the algorithm and the exploitation of good days is limited. Increasing the actionspace interval would improve the algorithm and the results would converge to the results of unconfined adaptive fractionation. In a potential application, choosing a large actionspace, i.e. a low minimum dose and a high maximum dose for each case should be pursued to maximize the improvement of adaptive fractionation, while the application of a minimum dose allows to reduce the risk of obtaining a significantly worse result than when using the reference treatment. The extension with the risk factor lead to similar results as the dose constraint with only small differences and can be considered as an additional option to reduce the probability of delivering an inferior treatment plan.

Compared to the research stated before by Lu et al.[3] and Chen et al.[4], the relative dose difference between the reference plan and adaptive fractionation is significantly lower, since both papers reported a decrease of OAR dose between 7.5%-22%/30% using adaptive fractionation. The average relative difference in tumor BED in this work is in the magnitude of a few percent. This large difference may originate from two key differences. In this analysis, the standard deviations of the probability distributions, which were motivated by observed patient data, were notably lower than in the mentioned papers, where standard deviations between 0.1 and 0.6 were used to model the distance between tumor and OAR. Furthermore, the fractionation schemes used in their research were not motivated by SBRT, but followed treatments with 40 fractions. Such treatments give more freedom to the algorithm to deliver lower doses on bad days and if probability updating is used, the probability distribution becomes more accurate than when using 6 sparing factors. Additionally, in both papers, no probability updating was conducted, but the probability distribution of the sparing factors was assumed to be known a priori which improves the quality of the treatment plans as well.

## 8 Outlook

Based on the evaluation of the patient data, adaptive fractionation at the MR-Linac could be performed with an expected average improvement of treatment plans. One big limitation of this work is the modeling of organ movement. All the geometrical information about the tumor and the OAR is reduced to one single number, discarding a large amount of information. Just the sparing factor definition based on different target volumes leads to different results. A more refined method to include organ motion into the algorithm may provide more possibilities to improve the quality of adaptive fractionation and the prediction of future patient geometries.

So far, the treatment optimization was based on the BED model and aiming on tumor dose escalation. A direct optimization of tumor control probability and minimization of the normal tissue complication probability gives the opportunity to pursue alternative objectives that could be of interest. The resulting optimal doses would not change if the same objectives are applied since both models are derived from the linear-quadratic model, but different objectives such as maximizing the TCP/NTCP ratio or other relevant objectives could be pursued.

In general, the algorithms can be extended with several functions. As stated before, the analyzed patients only had one dose limiting OAR. In practice, several organs can be in proximity of the tumor leading to a more complex dose computation, as several OAR limits must be met. An extension to multiple OARs can be conducted similar to the 3D algorithm, where more than one tissue BED is tracked. Additionally, the algorithms can be altered, such that the number of fractions are not fixed, but minimized as a new objective. Instead of determining the number of fractions at the beginning of treatment, a maximum amount of fractions can be set. The algorithm can then optimize the treatment such that the prescribed dose can be achieved in a minimum amount of fractions subject to the OAR limit. Such an objective can be valuable especially for the MR-Linac which has a workload.

The objectives and possibilities of adaptive fractionation are broad, depending on the goals that are relevant for clinical application. Further studies should aim on including more real patient data to assess the payoff of adaptive fractionation. In addition, a recalculation of dose distribution based on the optimal doses given by the optimization algorithm, i.e. not just scaling the dose to be delivered, but applying new prescribed tumor doses and OAR limits, may be an interesting subject of studies as the computed dose distributions largely depend on the defined tissues constraints.

## 9 Acknowledgements

I would first like to thank my supervisor Prof. Dr. Jan Unkelbach for giving me the opportunity to conduct my master project in his group at the University Hospital of Zurich in the Department of Radiation Oncology. His guidance and advice had a substantial contribution to the development of this thesis. I am also very grateful to Roman Ludwig who supported my project with weekly meetings alongside Prof. Dr. Jan Unkelbach and added very valuable inputs to the project.

I would also like to thank Dr. Riccardo Dal Bello and PD Dr. Stephanie Tanadini-Lang who guided and supported the acquisition of patient data and contributed essential clinical information.

A last acknowledgement goes to my family and my partner who supported me during all my studies.

## 10 Apendix

### 10.1 Code and patient data

The code for adaptive fractionation and all interfaces have been stored and made accessible in a github repository. The DVHs and the respective plots are also available in a separate folder.

[Repository link](#)

### 10.2 Tables and Figures

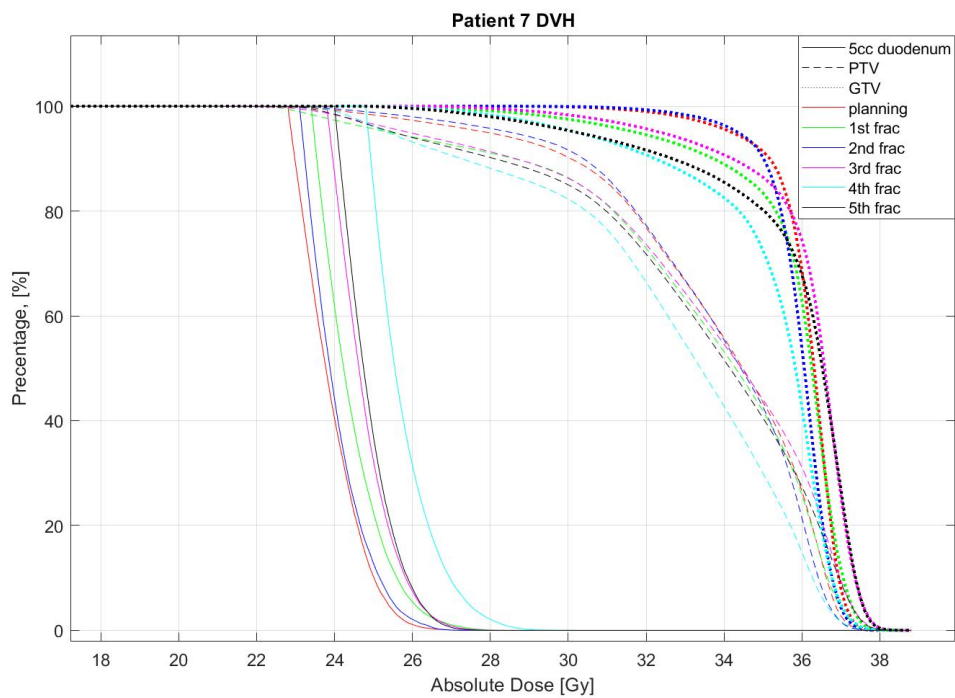


Figure 32: Dose volume histogram of patient 7. All relevant volumes are outlined with each fraction shown in a different color.



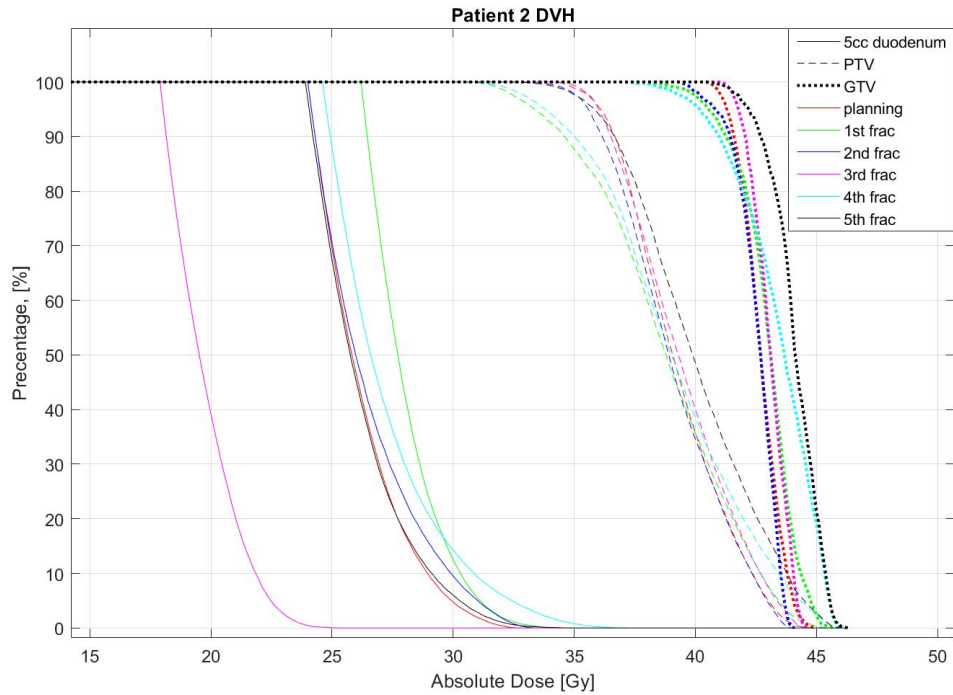


Figure 33: Dose volume histogram of patient 8. All relevant volumes are outlined with each fraction shown in a different color. Fraction three (magenta) stands out with an exceptionally low sparing factor.

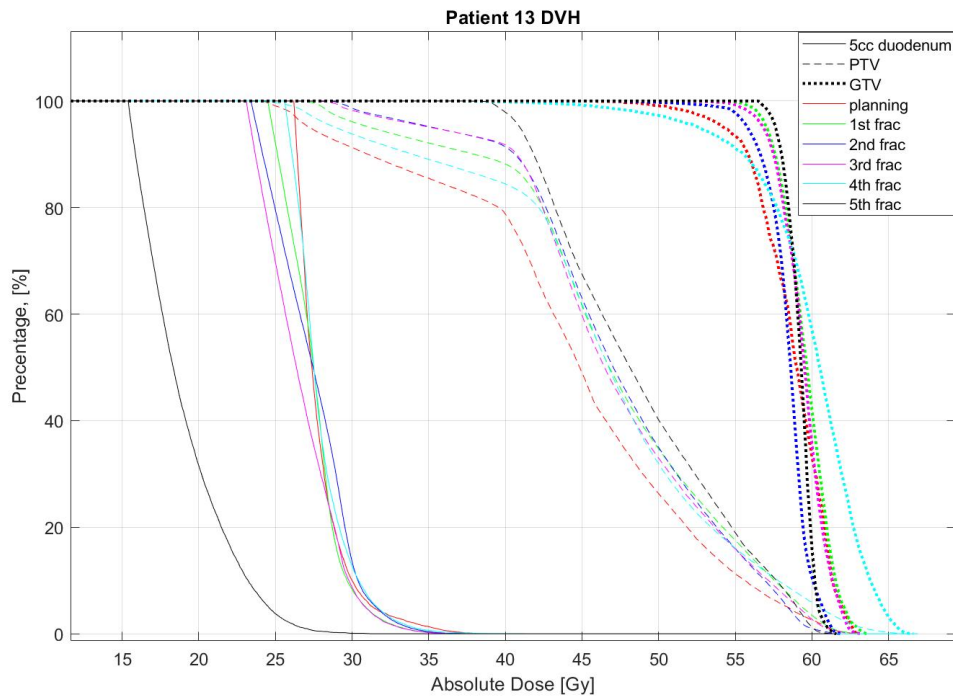


Figure 34: Dose volume histogram of patient 13. All relevant volumes are outlined with each fraction shown in a different color. Fraction five (black) stands out with an exceptionally low sparing factor.

patient	plan type	planning	fraction 1	fraction 2	fraction 3	fraction 4	fraction 5
1	sparing factor	0.99	0.95	0.98	0.96	1.02	1.01
	adaptive		7	5.7	6.4	5	6.35
	optimal		7.7	6	7.1	4.6	4.9
2	sparing factor	0.9	0.91	0.92	0.88	0.9	0.89
	adaptive		6.4	6.2	7.8	6.4	6.5
	optimal		6	5.6	7.9	6.6	7.2
3	sparing factor	0.73	0.67	0.78	0.53	0.89	0.73
	adaptive		12	4.1	20.7	0.3	2.17
	optimal		1.6	0.4	27.9	0	0.8
4	sparing factor	0.95	1	0.92	0.92	1.04	1.01
	adaptive		5.2	8	7	4	6.36
	optimal		4.7	8.7	8.7	3.7	4.4
5	sparing factor	0.94	0.83	0.97	1.13	1.05	0.96
	adaptive		9.9	4	2.4	4.9	8.64
	optimal		14.9	4	1.7	2.5	4.2
6	sparing factor	1.04	1.03	1.1	1.04	1.14	1.08
	adaptive		5.9	4.5	6.3	4.4	4.63
	optimal		7.3	4.6	6.7	3.8	5.2
7	sparing factor	0.88	0.99	0.87	0.98	1.04	1
	adaptive		4.4	9	4.5	4.3	7.95
	optimal		4.4	12.6	4.7	3.3	4.1
8	sparing factor	0.77	0.88	0.8	0.58	0.86	0.77
	adaptive		4.4	8.7	19.6	0.9	2.17
	optimal		0.49	1.1	25.3	0.6	1.4
9	sparing factor	1.04	1.02	0.99	1.02	1.08	1.05
	adaptive		6.1	6.7	5.5	4.6	6.1
	optimal		6	7.6	6	4.2	5
10	sparing factor	0.86	0.78	0.79	0.881	0.83	0.73
	adaptive		10.1	7.7	6.4	5.9	7.7
	optimal		6.8	6.1	5	4.4	14
11	sparing factor	1.01	0.83	0.8	0.82	0.85	0.91
	adaptive		11.4	8.6	5.7	4.3	4
	optimal		7.2	10.2	8	5.9	4
12	sparing factor	0.88	0.94	0.88	0.97	0.96	0.85
	adaptive		5.3	7.9	4.7	6	8.65
	optimal		4.8	7.9	3.9	4.2	10.9
13	sparing factor	1.06	0.92	0.84	0.82	1.01	0.53
	adaptive		9	9.6	6.8	2.4	7.27
	optimal		0	0.1	0.2	0	28.3
14	sparing factor	0.84	0.85	0.83	0.87	0.82	0.84
	adaptive		6.8	7.7	9.8	6	5.85
	optimal		5.2	6.3	11.4	7	5.7
15	sparing factor	0.78	0.72	0.74	0.83	0.79	0.79
	adaptive		10.7	7.7	4.1	7.1	8.7
	optimal		13.3	9.6	3.7	5.3	5.3
16	sparing factor	0.98	1.05	1.06	1.06	0.9	0.74
	adaptive		4.8	5.1	5.5	9.7	5.66
	optimal		1.3	1.3	1.3	3	18.8

Table 10: Delivered doses in adaptive fractionation and in the optimal plan.

---

## References

- [1] Fowler, J. F. Development of radiobiology for oncology—a personal view. *Physics in Medicine and Biology* **51** (2006).
- [2] Lajtha, L. G., Oliver, R. & Ellis, F. Rationalisation of Fractionation in Radiotherapy. *The British Journal of Radiology* **33**, 634–635 (1960). URL <http://www.birpublications.org/doi/10.1259/0007-1285-33-394-634>.
- [3] Lu, W., Chen, M., Chen, Q., Ruchala, K. & Olivera, G. Adaptive fractionation therapy: I. Basic concept and strategy. *Physics in Medicine and Biology* **53**, 5495–5511 (2008). URL <https://iopscience.iop.org/article/10.1088/0031-9155/53/19/015>.
- [4] Chen, M., Lu, W., Chen, Q., Ruchala, K. & Olivera, G. Adaptive fractionation therapy: II. Biological effective dose. *Physics in Medicine and Biology* **53**, 5513–5525 (2008). URL <https://iopscience.iop.org/article/10.1088/0031-9155/53/19/016>.
- [5] Guckenberger, M., Wilbert, J., Richter, A., Baier, K. & Flentje, M. Potential of Adaptive Radiotherapy to Escalate the Radiation Dose in Combined Radiochemotherapy for Locally Advanced Non–Small Cell Lung Cancer. *International Journal of Radiation Oncology\*Biography\*Physics* **79** (2011).
- [6] Wu, C., Jeraj, R., Olivera, G. H. & Mackie, T. R. Re-optimization in adaptive radiotherapy. *Physics in Medicine and Biology* **47**, 3181–3195 (2002). URL <https://iopscience.iop.org/article/10.1088/0031-9155/47/17/309>.
- [7] Joiner, M. C. & van der Kogel, A. *Basic Clinical Radiobiology, Fifth Edition* (Taylor & Francis, 2016). URL <https://books.google.ch/books?id=i91kNAEACAAJ>.
- [8] Barrett, A., Morris, S., Dobbs, J. & Roques, T. *Practical Radiotherapy Planning*. 1 (CRC Press, 2009). URL <https://www.taylorfrancis.com/books/9781444113112>.
- [9] Ahnesjö, A. & Aspradakis, M. M. Dose calculations for external photon beams in radiotherapy. *Physics in Medicine and Biology* **44** (1999).
- [10] Yang, Y. & Xing, L. Optimization of radiotherapy dose-time fractionation with consideration of tumor specific biology. *Medical Physics* **32**, 3666–3677 (2005).
- [11] Baskar, R., Dai, J., Wenlong, N., Yeo, R. & Yeoh, K.-W. Biological response of cancer cells to radiation treatment. *Frontiers in Molecular Biosciences* **1** (2014). URL <http://journal.frontiersin.org/article/10.3389/fmolb.2014.00024/abstract>.
- [12] Frometa-Castillo, T., Pyakuryal, A., Wals-Zurita, A. & Mesbahi, A. Biologically Effective Dose (BED) or Radiation Biological Effect (RBEf)? In *Recent Techniques and Applications in Ionizing Radiation Research*, 13 (IntechOpen, 2020). URL <http://dx.doi.org/10.1039/C7RA00172J%0Ahttps://www.intechopen.com/books/advanced-biometric-technologies/liveness-detection-in-biometrics%0Ahttp://dx.doi.org/10.1016/j.colsurfa.2011.12.014https://www.intechopen.com/books/recent-techniques-and-applications-in-ionizing-radiation-research/biologically-effective-dose-bed-or-radiation-biological-effect-rbef->.
- [13] van Leeuwen, C. M. *et al.* The alfa and beta of tumours: a review of parameters of the linear-quadratic model, derived from clinical radiotherapy studies. *Radiation Oncology* **13**, 96 (2018). URL <https://ro-journal.biomedcentral.com/articles/10.1186/s13014-018-1040-z>.
- [14] Sauer, R. *Strahlentherapie und Onkologie*. Elsevier (2010).

- 
- [15] Lo, S. S. *et al.* Stereotactic body radiation therapy: a novel treatment modality. *Nature Reviews Clinical Oncology* **7**, 44–54 (2010). URL <http://www.nature.com/articles/nrclinonc.2009.188>.
- [16] Ten Haken, R. K., Balter, J. M., Marsh, L. H., Robertson, J. M. & Lawrence, T. S. Potential benefits of eliminating planning target volume expansions for patient breathing in the treatment of liver tumors. *International Journal of Radiation Oncology\*Biological\*Physics* **38**, 613–617 (1997). URL <https://linkinghub.elsevier.com/retrieve/pii/S0360301697000096>.
- [17] Jaffray, D. A. Image-guided radiotherapy: From current concept to future perspectives. *Nature Reviews Clinical Oncology* **9**, 688–699 (2012).
- [18] Verellen, D. *et al.* Innovations in image-guided radiotherapy. *Nature Reviews Cancer* **7**, 949–960 (2007).
- [19] Hörner-Rieber, J. *et al.* MR-Guided Radiotherapy: The Perfect Partner for Immunotherapy? *Frontiers in Oncology* **10**, 1–8 (2021). URL <https://www.frontiersin.org/articles/10.3389/fonc.2020.615697/full>.
- [20] Klüter, S. Technical design and concept of a 0.35 T MR-Linac. *Clinical and Translational Radiation Oncology* **18**, 98–101 (2019). URL <https://linkinghub.elsevier.com/retrieve/pii/S2405630819300692>.
- [21] Abbas, H., Chang, B. & Chen, Z. J. Motion management in gastrointestinal cancers. *Journal of gastrointestinal oncology* **5**, 223–35 (2014). URL <http://www.ncbi.nlm.nih.gov/pubmed/24982771><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4074952>.
- [22] Ramakrishnan, J., Craft, D., Bortfeld, T. & Tsitsiklis, J. N. A dynamic programming approach to adaptive fractionation. *Physics in Medicine and Biology* **57**, 1203–1216 (2012). 1109.4524.
- [23] Jones, B., Dale, R. G., Deehan, C., Hopkins, K. I. & Morgan, D. A. L. The Role of Biologically Effective Dose (BED) in Clinical Oncology. *Clinical Oncology* **13**, 71–81 (2001). URL <https://www.sciencedirect.com/science/article/pii/S0936655501992210>.
- [24] Littman, M. Markov Decision Processes. *International Encyclopedia of the Social Behavioral Sciences* 9240–9242 (2001).
- [25] Garcia, F. & Rachelson, E. Markov Decision Processes. In *Markov Decision Processes in Artificial Intelligence*, 1–38 (John Wiley Sons, Inc., Hoboken, NJ USA, 2013). URL <https://onlinelibrary.wiley.com/doi/10.1002/9781118557426.ch1>.
- [26] Sutton, R. S. & Barto, A. G. *Reinforcement Learning: An Introduction* (A Bradford Book, Cambridge, MA, USA, 2018).
- [27] Murphy, K. P. Conjugate Bayesian analysis of the Gaussian distribution (2007).